www.rbmonline.com VOLUME 49 NO 1 JULY 2024

RBMO

RBMO

REPRODUCTIVE BIOMEDICINE ONLINE

Volume 49 number 1 July 2024

A journal covering the full range of research on human reproductive health and disease, from basic biology to the social implications of fertility care

RBMO

ELSEVIER

COUNTERCURRENT

Cryopreservation of ovarian tissue for fertility preservation in breast cancer patients: time to stop?

Kirsten Tryde Macklon[1], Michel De Vos[2]

ABSTRACT

Fertility preservation is currently offered to young women with breast cancer to increase their chances of motherhood after a potentially gonadotoxic treatment. Ovarian stimulation with oocyte vitrification and cryopreservation of ovarian tissue remain the most commonly used methods of choice. Whichever method is preferred is very much dependent on the practice and experience of the clinics, although for breast cancer in particular one method might be superior to the other. Cryopreservation of ovarian tissue is inevitably associated with the iatrogenic reduction of the ovarian reserve of a patient and should only be offered to women with a high risk of premature ovarian insufficiency following treatment. However, for younger breast cancer survivors, pregnancy and delivery rates are reassuringly high, even after chemotherapy. Despite its widespread use, few women come back to make use of their cryopreserved tissue. It is argued here that cryopreservation of ovarian tissue is not an appropriate option for breast cancer patients and discuss the reasons for this opinion.

METHODS TO PRESERVE FERTILITY IN BREAST CANCER PATIENTS

Fertility preservation before the initiation of chemotherapy has become embedded in the care path of young, premenopausal women with breast cancer. Ovarian stimulation to obtain a set of oocytes that can be vitrified has been recommended as the fertility preservation strategy of first choice. Although ovarian stimulation can be started at any time of the menstrual cycle, this fertility preservation approach may delay the start of the cancer treatment by approximately 10–14 days (*Greer et al., 2021*). However, this delay, if any, was not observed in an observational study (*Letourneau et al., 2017*), and should not jeopardize the effectiveness of cancer treatment (*de Melo Gagliato et al., 2020*). In the case of hormone receptor-positive breast cancer, co-administration of aromatase inhibitors such as letrozole can be considered to mitigate supraphysiological oestradiol concentrations. The results in terms of pregnancy rates and live birth rates using cryopreserved oocytes now equal those obtained from the cryopreservation and thawing of embryos, at least in the most experienced laboratories.

While vitrification of oocytes following ovarian stimulation is first choice according to most guidelines, including the European Society of Human Reproduction and Embryology guideline on female fertility preservation (*Anderson et al., 2020*), alternative methods such as ovarian tissue cryopreservation (OTC), co-treatment with a gonadotrophin-releasing hormone (GnRH) agonist or in-vitro maturation of immature oocytes (IVM) are also offered to this patient group; this is for various reasons, including the reluctance of some oncologists to allow ovarian stimulation in hormone receptor-positive breast cancer patients (*Biglia et al., 2015*). OTC requires the removal of ovarian tissue either as a whole ovary or as ovarian cortical biopsies, typically performed during laparoscopy and under general anaesthesia. When the patient wants to make use of her tissue, another surgical procedure is necessary, as autotransplantation is still the only way possible to use the eggs in the tissue.

FERTILITY AFTER BREAST CANCER

There has been a longstanding belief that the chemotherapy protocols used for the treatment of early-stage breast cancer are very detrimental to ovarian function and that breast cancer patients face a high risk of infertility after treatment. Thus, the general perception of infertility as one of the major concerns among young breast cancer patients, and the emphasis on the importance of fertility counselling and preservation to enhance patients' adherence to cancer treatment, have

[1] Fertility Department, University Hospital of Copenhagen, Rigshospitalet, Copenhagen, Denmark
[2] Brussels IVF, Universitair Ziekenhuis Brussel, Vrije Universiteit, Brussels, Belgium

E-mail address: https://doi.org/10.1016/j.rbmo.2024.103939 1472-6483/© 2024 Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS
Breast cancer
Cryopreservation
Female
Fertility preservation
Ovarian
Ovary

made breast cancer the most common indication for fertility preservation for adult patients. Indeed, a recent systematic review and meta-analysis showed that breast cancer survivors had a 60% reduced likelihood of having a subsequent pregnancy compared with the background population, only surpassed by cervical cancer survivors (*Lambertini et al., 2021*). These data were backed up by a Swedish study that found that only 8.7% of breast cancer patients not exposed to fertility preservation experienced a live birth within a mean follow-up period of 4.6 years after treatment (*Marklund et al., 2020*), as opposed to 22.8% of those exposed to fertility preservation.

However, other studies have reported much more favourable estimates of ovarian function and fertility after the treatment of breast cancer. Schmidt and colleagues looked at fertility in 143 patients who had received OTC of an entire ovary before cancer treatment, of whom 54 had a diagnosis of breast cancer (*Schmidt et al., 2013*). After treatment 22 had a pregnancy wish, 16 (73%) obtained a pregnancy and 14 (64%) delivered. All the pregnancies had been conceived naturally, and none of the women had made use of their cryopreserved tissue. The women had a mean age at OTC of 30.2 years and a mean age at follow-up of 35.5 years. The majority became pregnant within the first 6 months of trying. Equally reassuring, a Dutch study reported on their experience with fertility preservation counselling in 118 young women aged 19–40 years with breast cancer and found a 5-year ovarian function rate of 92% based on menstrual cycles and laboratory values in the premenopausal range (*Vriens et al., 2020*).

Very recently, results from the POSITIVE trial demonstrated a high pregnancy rate of 74% and a delivery rate of 63.8% in a group of 516 women with a breast cancer diagnosis (*Partridge et al., 2023*), with menstrual cycles returning within 6 months of pausing endocrine therapy in most patients who were amenorrhoeic at trial entry. Not surprisingly, patient age and the use of cryopreserved oocytes/embryos were correlated with the chance of pregnancy (personal communication, Hatem A. Azim Jr., Cairo Cure Oncology Center). Together, these studies indicate that we may have been too negative in estimating the fertility potential in women following the treatment of breast cancer.

So why this discrepancy between the studies? One reason could be that most of the studies reporting on low fertility rates and high premature ovarian insufficiency (POI) rates after breast cancer included all premenopausal women, i.e. also including women in their 40s who already have challenges to their fertility even before chemotherapy due to the age-related natural decline in the number of healthy follicles in the ovaries. Another important factor is the length of the follow-up time. It has been found that there is a significant period of recovery of ovarian function after treatment with chemotherapy, with a non-linear decrease in anti-Müllerian hormone concentration seen in most patients during chemotherapy followed by an increase over the next 2–4 years (*Zhou et al., 2022*). So, if the follow-up period is too short, a lower estimate of ovarian recovery may be seen than if a longer follow-up period is allowed.

INTERRUPTION OF ENDOCRINE THERAPY FOR OESTROGEN RECEPTOR-POSITIVE BREAST CANCER

For women with oestrogen receptor-positive breast cancer the recommendations are 5–10 years of endocrine therapy during which time pregnancy is contraindicated. Endocrine therapy consists of either a selective oestrogen receptor modulator such as tamoxifen, or an aromatase inhibitor such as anastrozole. This has been shown to improve the survival rate.

However, lately this notion has been challenged with results from the POSITIVE trial in which women with early-stage hormone receptor-positive breast cancer and a pregnancy wish were allowed to discontinue their adjuvant endocrine therapy in order to attempt a pregnancy (*Partridge et al., 2023*). The results were very encouraging, as a temporary interruption of the endocrine therapy did not give rise to a greater short-term risk of breast cancer events. Additionally, in this study in which the women were prospectively followed, the 3-year cumulative incidences of breast cancer events and distant recurrences were the same as in the external control cohort. This is a very important finding for young women with breast cancer, as 10 years of endocrine therapy for many will mean that by the time they are finally allowed to become pregnant their oocyte quality will have declined and their ovaries will be subject to the natural age-related depletion of follicles on top of the chemotherapy-induced depletion, making it difficult to conceive.

LOW RETURN RATES AFTER OTC

There is a lack of longitudinal studies investigating the utilization of cryopreserved material after fertility preservation in breast cancer patients, and a crucial lack of information about reproductive choices and outcomes in women who have not come back to use their gametes or ovarian tissue. Based on the available literature, the numbers of breast cancer survivors who eventually embark on ovarian tissue transplantation (OTT) are consistently small. Colmorn and colleagues reported 19 Danish breast cancer patients who underwent OTT between 2003 and 2021 (*Colmorn et al., 2022*), Lotz and co-workers reported 82 breast cancer patients who has OTT between 2007 and 2019 within the FertiPROTEKT network (*Lotz et al., 2022*), and the systematic review by Khattak and collaborators encompassed 87 studies reporting outcomes in 54 breast cancer survivors who underwent OTT (*Khattak et al., 2022*).

The small number of patients returning for OTT may be related to the low risk of POI after breast cancer treatment; indeed, data from 225 patients who had OTC in a single centre in Belgium before gonadotoxic therapy revealed that POI was found in only 11% of breast cancer survivors compared with 34.5% of those with haematological diseases after a follow-up period of at least 1 year (*Imbert et al., 2014*). Nevertheless, low return rates have also been observed in follow-up studies after vitrification of oocytes. There is a clear need for an international registry for the follow-up of patients who have undergone fertility preservation generally, and for breast cancer patients specifically, to help practitioners counsel breast cancer patients about fertility preservation strategies.

THE ROLE OF *BRCA1* AND *BRCA2* MUTATIONS

A non-negligible concern that arises when young women with a diagnosis of breast cancer embark on OTC is the identification of a novel deleterious mutation in a cancer-predisposing gene in a substantial proportion of these women, usually several months after their diagnosis.

In these women, harvesting a maximal number of oocytes before the start of the cancer treatment (and before they know their genetic status) should boost their chances of a transferrable embryo when they return to the fertility clinic for assisted reproductive technology (ART) with preimplantation genetic testing (PGT) to avoid the transmission of the mutation. If these women had opted for OTC instead, they would have faced a theoretical risk of a low ovarian response as a result of the combination of surgical damage to the ovary and chemotherapy, which could reduce their chance of a successful PGT treatment. Moreover, OTT in women who carry a mutation in a *BRCA* gene and for whom prophylactic bilateral salpingo-oophorectomy is recommended evokes even more questions on the role of OTC in breast cancer patients.

CONCLUSIONS

With this paper we challenge the practice of offering OTC to breast cancer patients. While we acknowledge that offering oocyte vitrification should enhance psychological well-being and adherence to treatment in these patients, and that there may be a case for co-treatment with a GnRH agonist or even IVM in young women with early-stage breast cancer, we believe that the time has come to stop offering OTC as a means of fertility preservation in breast cancer patients. First, there has probably been a tendency to underestimate the fertility potential after chemotherapy, at least in women under 40 years of age, as high pregnancy and delivery rates are seen in breast cancer survivors in this age group, with and without the use of ART. Secondly, the low return rates after OTC in these patients mean that freezers in clinics around the world are being filled with 'orphaned' ovaries that will never be used.

DATA AVAILABILITY

No data was used for the research described in the article.

REFERENCES

ESHRE Guideline Group on Female Fertility Preservation, Anderson, RA., Amant, F., Braat, D., D'Angelo, A., Chuva de Sousa Lopes, SM., Demeestere, I., Dwek, S., Frith, L., Lambertini, M., Maslin, C., Moura-Ramos, M., Nogueira, D., Rodriguez-Wallberg, K., Vermeulen, N., 2020. ESHRE guideline: female fertility preservation. Hum Reprod Open 2020 (4), hoaa052. https://doi.org/10.1093/hropen/hoaa052 PMID: 33225079; PMCID: PMC7666361.

Biglia, N., Torrisi, R., D'Alonzo, M., Pisanelli, GC., Rota, S., Peccatori, FA, 2015. Attitudes on fertility issues in breast cancer patients: an Italian survey. Gynecol Endocrinol 31, 458–464.

Colmorn, LB., Pedersen, AT., Larsen, EC., Hansen, AS., Rosendahl, M., Yding Andersen, C., Kristensen, SG., Macklon, KT., 2022. Reproductive and Endocrine Outcomes in a Cohort of Danish Women Following Auto-Transplantation of Frozen/Thawed Ovarian Tissue from a Single Center. Cancers 14, 5873. https://doi.org/10.3390/cancers14235873.

De Melo Gagliato, D., Lei, X., Giordano, SH., Valero, V., Barcenas, CH., Hortobagyi, GN., Chavez-MacGregor, M., 2020. Impact of delayed neoadjuvant systemic chemotherapy on overall survival among patients with breast cancer. Oncologist 25, 749–757.

Greer, AC., Lanes, A., Poorvu, PD., Kennedy, P., Thomas, AM., Partridge, AH., Ginsburg, ES., 2021. The impact of fertility preservation on the timing of breast cancer treatment, recurrence, and survival. Cancer 127, 3872–3880.

Imbert, R., Moffa, F., Tsepelidis, S., Simon, P., Delbaere, A., Devreker, F., Dechene, J., Ferster, A., Veys, I., Fasrez, M., Englert, Y., Demeestere, I., 2014. Safety and usefulness of cryopreservation of ovarian tissue to preserve fertility: a 12-year retrospective analysis. Hum Reprod 29, 1931–1940.

Khattak, H., Malhas, R., Craciunas, L., Afifi, Y., Amorim, CA., Fishel, S., Silber, S., Gook, D., Demeestere, I., Bystrova, O., Lisyanskaya, A., Manikhas, G., Lotz, L., Dittrich, R., Colmorn, LB., Macklon, KT., Hjorth, IMD., Kristensen, SG., Gallos, I., Coomarasamy, A., 2022. Fresh and cryopreserved ovarian tissue transplantation for preserving reproductive and endocrine function: a systematic review and individual patient data meta-analysis. Hum Reprod Update 28 (3), 400–416.

Lambertini, M., Blondeaux, E., Bruzzone, M., Perachino, M., Anderson, RA., de Azambuja, E., Poorvu, PD., Kim, HJ., Villarreal-Garza, C., Pistilli, B., Vaz-Luis, I., Saura, C., Ruddy, KJ., Franzoi, MA., Sertoli, C., Ceppi, M., Azim, Jr, HA., Amant, F., Demeestere, I., Del Mastro, L., Partridge, AH., Pagani, O., Peccatori, FA, 2021 Oct 10. Pregnancy After Breast Cancer: A Systematic Review and Meta-Analysis. J Clin Oncol 39 (29), 3293–3305.

Letourneau, JM., Sinha, N., Wald, K., Harris, E., Quinn, M., Imbar, T., Mok-Lin, E., Chien, AJ., Rosen, M., 2017. Random start ovarian stimulation for fertility preservation appears unlikely to delay initiation of neoadjuvant chemotherapy for breast cancer. Hum Reprod 32, 2123-2129.55.

Lotz, L., Bender-Liebenthron, J., Dittrich, R., Häberle, L., Beckmann, MW., Germayer, A., Korell, M., Sänger, N., Kruesell, JS., von Wolff, M., FertiPROTEKT (Transplantation group), 2022. Determinants of transplantation success with cryopreserved ovarian tissue: data from 196 women of the FertiPROTEKT network. Hum Reprod 37, 2787–2796.

Marklund, A., Lundberg, FE., Eloranta, S., Hedayati, E., Petterson, K., Rodriguez-Wallberg, KA., 2020. Reproductive Outcomes After Breast Cancer in Women With vs. Without Fertility Preservation. JAMA Oncol 7, 86–91.

Partridge, AH., Niman, SM., Ruggeri, M., Peccatori, FA., Azim, Jr, HA., Colleoni, M., Saura, C., Shimizu, C., Sætersdal, AB., Kroep, JR., Mailliez, A., Warner, E., Borges, VF., Amant, F., Gombos, A., Kataoka, A., Rousset-Jablonski, C., Borstnar, S., Takei, J., Lee, JE., Walshe, JM., Ruíz-Borrego, M., Moore, HCF., Saunders, C., Bjelic-Radisic, V., Susnjar, S., Cardoso, F., Smith, KL., Ferreiro, T., Ribi, K., Ruddy, K., Kammler, R., El-Abed, S., Viale, G., Piccart, M., Korde, LA., Goldhirsch, A., Gelber, RD., Pagani, O., International Breast Cancer Study Group; POSITIVE Trial Collaborators, 2023. Interrupting Endocrine Therapy to Attempt Pregnancy after Breast Cancer. N Engl J Med 388 (18), 1645–1656.

Schmidt, KT., Nyboe Andersen, A., Greve, T., Ernst, E., Loft, A., Yding Andersen, C., 2013. Fertility in cancer patients after cryopreservation of one ovary. Reprod Biomed Online 26, 272–279.

Vriens, IJH., Ter Welle-Butalid, EM., de Boer, M., de Die-Smulders, CEM., Derhaag, JG., Geurts, SME., van Hellemond, IEG., Luiten, EJT., Dercksen, MW., Lemaire, BMD., van Haaren, ERM., Vriens, BEPJ., van de Wouw, AJ., van Riel, AMGH., Janssen-Engelen, SLE., van de Poel, MHW., Schepers-van der Sterren, EEM., van Golde, RJT., Tjan-Heijnen, VCG, 2020. Preserving fertility in young women undergoing chemotherapy for early breast cancer; the Maastricht experience. Breast Cancer Res Treat 181 (1), 77–86.

Zhou, B., Kwan, B., Desai, MJ., Nalawade, V., Ruddy, KJ., Nathan, PC., Henk, HJ., Murphy, JD., Whitcomb, BW., Su, HI., 2022. Long-term Anti-Müllerian hormone patterns differ by cancer treatment exposures in young breast cancer survivors. Fertil Steril 117, 1047–1056.

Received 9 February 2024; received in revised form 29 February 2024; accepted 29 February 2024.

RBMO

ELSEVIER

COUNTERCURRENT

Endometriosis, staging, infertility and assisted reproductive technology: time for a rethink

Baris Ata[1,2,*], Edgardo Somigliana[3,4]

ABSTRACT

How endometriosis causes infertility, with the exception of tubal dysfunction caused by adhesions, is unclear. The inflammatory milieu in the pelvis and impaired receptivity of the eutopic endometrium are considered to be possible factors. Anatomical staging systems fail to predict the fertility status of endometriosis patients. Data from assisted reproductive technology cycles consistently suggest that oocytes from patients with endometriosis have a normal potential to develop into euploid blastocysts. Moreover, oocyte or embryo recipients with endometriosis seem to have similar or slightly lower pregnancy and live birth rates compared with recipients without endometriosis, suggesting that eutopic endometrium is not or is only minimally affected, which may be caused by undiagnosed adenomyosis. In-vivo observations from women with endometriomas provide evidence against a detrimental effect of endometriomas on oocytes. Combined with the absence of an obvious improvement in fertility following the surgical destruction or excision of peritoneal endometriosis or from temporary medical suppression of the disease and the associated inflammation, the available evidence makes endometriosis-associated infertility questionable in the absence of tubal dysfunction caused by adhesions. It is likely that no anatomical staging will correlate with fertility beyond assessing tubal function. In patients with endometriosis assisted reproductive technology is as effective as for other indications.

BACKGROUND

Endometriosis is an oestrogen-dependent chronic inflammatory fibrotic disease that is characterized by the presence of endometriotic foci outside the uterine cavity. The disease most likely occurs through a seeding of live endometrial cells into the pelvis by retrograde menstruation. While it is most often observed on the ovaries and pelvic peritoneum, ectopic foci can exist on other pelvic and extrapelvic organs, at times growing deeper beneath peritoneal surfaces.

Multiple staging systems have been developed to unify the description of the extent of spread in the pelvis and beyond. The purpose of staging is to define homogenous groups of patients who could be expected to experience similar symptoms and prognoses with and without treatment. Such a classification is expected to guide patient counselling and endometriosis research by enabling the formation of similar study populations to compare the effectiveness of various treatments with each other or with expectant management in terms of outcomes, which often include but are not limited to an improvement in pain or fertility. However, while staging correlates with pain, it does not correlate with fertility.

HOW DOES ENDOMETRIOSIS CAUSE INFERTILITY?

Endometriosis has been long recognized as a cause of infertility. Indeed, infertility is regarded as one of the two cardinal symptoms, along with chronic pelvic pain. While it is easily comprehensible that endometriosis causes tubal factor infertility through adhesions that directly block one or both Fallopian tubes, disrupt tubo-ovarian relationships or impede tubal access to the pouch of Douglas where follicular fluid containing an oocyte is often drained to, it is unclear how the disease affects fertility in the absence of tubal factor (*Somigliana et al., 2017*).

Let us concisely review how else endometriosis can impact female fertility. First and foremost, endometriosis patients are ovulatory in the absence of another pathology. The disease occurs by the retrograde seeding of endometrial cells to the pelvis, which is dependent on menstruation, a function of ovulation. Moreover, endometriosis is stimulated by oestrogen produced during folliculogenesis. Clearly oligo-anovulation does not contribute to endometriosis-associated infertility.

It has been long considered that endometriosis, particularly ovarian endometriosis, affects 'oocyte quality', a vague term in the context. However,

[1] Department of Obstetrics and Gynecology, Koç University School of Medicine, Istanbul, Turkey.
[2] ART Fertility Clinics, Dubai, United Arab Emirates.
[3] ART Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.
[4] Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy.

*Corresponding author. E-mail address: barisata@ku.edu.tr (B. Ata). https://doi.org/10.1016/j.rbmo.2024.103943 1472-6483/

Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS
Adhesions
Assisted reproductive technology
Endometriosis
Endometrial receptivity
Euploidy
Infertility

recent studies conducted on patients undergoing assisted reproductive technology (ART) cycles consistently demonstrate that oocytes retrieved from patients with endometriosis and endometriomas have a similar potential to be fertilized, develop into blastocysts and yield euploid blastocysts as is seen in various control groups undergoing ART without a diagnosis of endometriosis (*Ata and Telek, 2021*; *Somigliana et al., 2023*). The consistency of these findings from observational studies by different groups from multiple countries is striking and can be regarded as high-quality evidence.

Moreover, a landmark study comparing natural pregnancy rates between ovulation from an endometrioma-containing ovary and the contralateral ovary without an endometrioma in women with unilateral endometriomas also reported similar natural pregnancy rates, also suggesting the absence of a negative effect of endometriomas on the in-vivo potential of oocytes (*Leone Roberti Maggiore et al., 2015*). Overall, it seems that oocytes are spared from a theoretical adverse effect of endometriosis on their reproductive potential, at least under in-vitro laboratory conditions.

Endometriosis, particularly endometriomas and their surgical removal, affects ovarian reserve (*Yilmaz et al., 2019*). However, a large body of evidence unequivocally shows that, if the ovarian reserve is sufficient to ensure regular ovulation, it is almost irrelevant to natural conception (*Hvidman et al., 2016*; *Steiner et al., 2017*). The damage to the ovarian reserve can affect the ART success per stimulation cycle (*Ata and Telek, 2021*; *Somigliana et al., 2023*), but not natural conception.

Whether eutopic endometrium is less receptive in endometriosis patients is another controversy, yet data from observational studies suggest that euploid blastocysts enjoy similar implantation and live birth rates in patients with and without endometriosis (*Ata and Telek, 2021*; *Somigliana et al., 2023*). Moreover, pregnancy and live births following embryo transfer with or without preimplantation genetic testing for aneuploidies also seem similar in large registry studies. A recent meta-analysis showed that implantation and live birth rates in oocyte-donation cycles is not or is only slightly reduced in recipients with endometriosis (*Paffoni et al., 2024*).

Once adenomyosis, which is a close but separate entity and seems to affect implantation and pregnancy outcomes, is ruled out, endometriosis per se may have no or perhaps a minimal impact on endometrial receptivity. The observed minimal impact may be brought about by adenomyosis, which often coexists with endometriosis but may have been underdiagnosed. It is noteworthy that none of the current staging systems for endometriosis currently considers endometrial receptivity. How an anatomical staging system can include endometrial receptivity is unclear, since it is not visually assessable.

A potential detrimental effect of the inflammatory pelvic milieu on oocytes cannot be assessed with IVF studies since the oocytes are aspirated prior to exposure. Thus, endometriosis-related pelvic inflammation could be the culprit behind impaired natural fertility. In theory, pelvic inflammation could impair sperm function in the distal tube, fertilization or early embryo development. However, whether endometriosis of extragenital structures, such as the pelvic peritoneum, bladder or rectosigmoid colon, can affect fertility, except for possibly decreasing the frequency of intercourse due to pain, is uncertain. Notably, if an inflammatory detrimental mechanism does exist, one could assume that the resection of endometriosis would restore or improve natural fertility.

However, this does not seem to be the case for treatment of peritoneal endometriosis. A few studies, and probably more meta-analyses than studies themselves, report contradictory results depending on the inclusion/exclusion criteria. A clear and relevant benefit of the ablation or excision of peritoneal endometriosis has not been demonstrated. It seems that the wider academic and clinical community also has a similar opinion since routine laparoscopy to diagnose and treat such lesions to promote fertility is regarded as poorly effective and therefore a diagnostic laparoscopy is not recommended in the recent guidelines or not commonly practised, as in the past (*Becker et al., 2022*).

It should also be noted that a temporary suppression of endometriosis, and hence the associated inflammation, by gonadotrophin-releasing hormone analogues or oestroprogestins, does not seem to improve natural fertility or ART outcomes either, again speaking against a negative effect of such lesions on reproductive process (*Becker et al., 2022*). Of the utmost relevance here is that surgery can properly remove endometriotic lesions but is poorly effective for adhesions; in most cases, adhesions reform rapidly after surgery.

It is difficult to assess an isolated effect of deep endometriosis with or without bowel involvement on fertility since in most cases widespread adhesions are also present and distort the pelvic anatomy, and hence tubal function. Studies investigating whether the surgical treatment of bowel endometriosis improves fertility or ART outcomes are few and often limited by a retrospective design and lack of proper controls.

After decades of research, whether endometriosis causes infertility except for a tubal factor caused by adhesions seems controversial, and building evidence, particularly in the ART setting, suggests otherwise (*Ata and Telek, 2021*; *Somigliana et al., 2023*). One can argue that fertilization and early embryo development occur in the Fallopian tube, as opposed to what occurs in the ART setting, and perhaps it is the inflammatory environment that diminishes reproductive potential *in vivo*. However, it seems that temporary endocrine suppression of the disease improves neither the chance of a natural pregnancy nor ovulations from an ovary harbouring an endometrioma, which can be assumed to be on the side of a more affected Fallopian tube and have lower pregnancy and live birth rates than ovulations from the contralateral, relatively healthy side (*Leone Roberti Maggiore et al., 2015*).

In our opinion, endometriosis per se is unlikely to be a major aetiological factor for infertility, and this may be the reason for the lack of an association between staging focusing on the anatomical spread of the lesions and the patients' reproductive potential. Indeed, the only sort of staging system that correlates with natural fertility or the success of fertility-promoting treatments is the Endometriosis Fertility Index, which focuses on the functional status of the tubes and fimbriae, disregarding peritoneal and other extragenital endometriosis (*Adamson and Pasta, 2010*).

CONCLUSIONS

Except for tubal blockage or dysfunction, endometriosis per se seems to have minimal if any effect on fertility. Hence, any anatomical staging system will almost invariably fail to correlate with or predict the fertility status of a patient with endometriosis, except for tubal function. More interestingly, this may be the time to reconsider endometriosis as a detrimental diagnosis in terms of IVF success. It is not. The procedure overcomes the established negative effects on tubal function and is therefore highly successful provided that the ovarian reserve is preserved.

DATA AVAILABILITY

No data was used for the research described in the article.

REFERENCES

Adamson, G.D., Pasta, D.J., 2010. Endometriosis fertility index: the new, validated endometriosis staging system. Fertil. Steril. 94 (5), 1609–1615. https://doi.org/10.1016/j.fertnstert.2009.09.035 OctPMID: 19931076.

Ata, B., Telek, S.B., 2021. Assisted reproductive technology for women with endometriosis, a clinically oriented review. Curr. Opin. Obstet. Gynecol. 33 (3), 225–231. https://doi.org/10.1097/GCO.0000000000000710 Jun 1PMID: 33769421.

Becker, C.M., Bokor, A., Heikinheimo, O., Horne, A., Jansen, F., Kiesel, L., King, K., Kvaskoff, M., Nap, A., Petersen, K., Saridogan, E., Tomassetti, C., van Hanegem, N., Vulliemoz, N., Vermeulen, N., ESHRE Endometriosis Guideline Group, 2022. ESHRE guideline: endometriosis. Hum Reprod Open 2022 (2), hoac009. https://doi.org/10.1093/hropen/hoac009 Feb 26PMID: 35350465; PMCID: PMC8951218.

Hvidman, H.W., Bentzen, J.G., Thuesen, L.L., Lauritsen, M.P., Forman, J.L., Loft, A., Pinborg, A., Nyboe Andersen, A., 2016. Infertile women below the age of 40 have similar anti-Müllerian hormone levels and antral follicle count compared with women of the same age with no history of infertility. Hum Reprod 31 (5), 1034–1045. https://doi.org/10.1093/humrep/dew032 MayEpub 2016 Mar 9. PMID: 26965431.

Leone Roberti Maggiore, U., Scala, C., Venturini, P.L., Remorgida, V., Ferrero, S., 2015. Endometriotic ovarian cysts do not negatively affect the rate of spontaneous ovulation. Hum. Reprod. 30 (2), 299–307. https://doi.org/10.1093/humrep/deu308 FebEpub 2014 Nov 28. PMID: 25432923.

Paffoni, A., Casalechi, M., De Ziegler, D., Cicinelli, E., Somigliana, E., Viganò, P., Vitagliano, A., 2024. Live Birth After Oocyte Donation In Vitro Fertilization Cycles in Women With Endometriosis: A Systematic Review and Meta-Analysis. JAMA Netw. Open 7 (1), e2354249. https://doi.org/10.1001/jamanetworkopen.2023.54249 Jan 2PMID: 38294811; PMCID: PMC10831577.

Somigliana, E., Li Piani, L., Paffoni, A., Salmeri, N., Orsi, M., Benaglia, L., Vercellini, P., Vigano', P., 2023. Endometriosis and IVF treatment outcomes: unpacking the process. Reprod. Biol. Endocrinol 21 (1), 107. https://doi.org/10.1186/s12958-023-01157-8 Nov 7PMID: 37936154; PMCID: PMC10629090.

Somigliana, E., Vigano, P., Benaglia, L., Busnelli, A., Berlanda, N., Vercellini, P., 2017. Management of Endometriosis in the Infertile Patient. Semin Reprod Med 35 (1), 31–37. https://doi.org/10.1055/s-0036-1597125 JanEpub 2016 Dec 7. PMID: 27926972.

Steiner, A.Z., Pritchard, D., Stanczyk, F.Z., Kesner, J.S., Meadows, J.W., Herring, A.H., Baird, D.D., 2017. Association Between Biomarkers of Ovarian Reserve and Infertility Among Older Women of Reproductive Age. JAMA 318 (14), 1367–1376. https://doi.org/10.1001/jama.2017.14588 Oct 10PMID: 29049585; PMCID: PMC5744252.

Yılmaz Hanege, B., Güler Çekıç, S., Ata, B., 2019. Endometrioma and ovarian reserve: effects of endometriomata per se and its surgical treatment on the ovarian reserve. Facts Views Vis. Obgynk. 11 (2), 151–157 JunPMID: 31824636; PMCID: PMC6897522.

Received 20 February 2024; accepted 29 February 2024.

RBMO

COMMENTARY

The metrics maze in science: navigating academic evaluation without journalistic pressures

Carlo Alviggi[1], Rossella E. Nappi[2,3], Antonio La Marca[4], Filippo Maria Ubaldi[5], Alberto Vaiarelli[5,*]

ABSTRACT

In recent years a troubling trend has emerged in the medical research field, notably in reproductive medicine, manifesting an increased emphasis on quantity over quality in articles published.

The pressure to collect copious publication records risks compromising meticulous expertise and impactful contributions. This tendency is exemplified by the rise of 'hyper-prolific researchers' publishing at an extraordinary rate (i.e. every 5 days), prompting a deeper analysis of the reasons underlying this behaviour. Prioritizing rapid publication over Galileo Galilei's systematic scientific principles may lead to a superficial approach driven by quantitative targets. Thus, the overreliance on metrics to facilitate academic careers has shifted the focus to numerical quantification rather than the real scientific contribution, raising concerns about the effectiveness of the evaluation systems. The Hamletian question is: are we scientist or journalist? Addressing these issues could necessitate a crucial re-evaluation of the assessment criteria, emphasizing a balance between quantity and quality to foster an academic environment that values meaningful contributions and innovation.

In the dynamic landscape of academic research, the evaluation system has undergone a profound transformation, primarily driven by various metrics and indicators used to assess the quality of research conducted by universities and researchers. The advent of quantitative parameters and median indicators has provided an objective measure of researchers' scientific productivity, extending beyond universities and enabling a shift from subjective to objective evaluation. This pivotal shift has commendably safeguarded the recruitment system from nepotism and facilitated meritocratic access to academic careers for highly qualified individuals, even in private contexts.

However, in certain countries, notably in Italy, a concerning trend among young researchers has emerged. This trend is characterized by a prioritization of quantity over quality in publications. The pressure to accumulate a prolific publication record has intensified, often at the expense of the meticulous craftsmanship that distinguishes impactful contributions from mere statistical entries (*Conroy, 2024*). Indeed, some researchers, solely devoted to science and spending their entire lives confined in laboratories scattered across the world, are capable of publishing an article every 5 days (*Ioannidis et al., 2018*). While this may serve as an inspiring example for our young students and colleagues, the noteworthy increase in the number of these hyper-prolific researchers, quadrupling over the last decade, warrants a more thorough analysis of the reasons behind this surge in scientific research. In this regard, curiosity appears to be overshadowed by the necessity of publication, where the focus shifts from the aim of discovering something to the act of publishing itself.

Contrary to the foundational principles taught by Galileo Galilei, emphasizing systematic observation, hypothesis formulation, experimentation and the sharing of results with the scientific community, current trends favour rapid publication over the time-intensive scientific method. Young researchers, driven by the pursuit of meeting quantitative targets, may compromise the quality of their research. This shift could encourage a superficial approach, prioritizing meeting deadlines (as reporters must) rather than conducting a thorough and ground-breaking investigations (like investigative journalists).

Furthermore, the pressure to publish frequently may disproportionately affect early-career researchers, pushing them towards strategic collaborations in too many diverse areas. Moreover, such a 'hit and run strategy' makes it difficult to

[1] Department of Public Health, University of Naples Federico II, Naples, Italy.
[2] Department of Clinical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy.
[3] Research Center for Reproductive Medicine and Gynecological Endocrinology – Menopause Unit, IRCCS S. Matteo Foundation, Pavia, Italy.
[4] Department of Medical and Surgical Sciences of the Mother, Children and Adults, University of Modena and Reggio Emilia, Modena, Italy.
[5] IVIRMA Global Research Alliance, GENERA, Clinica Valle Giulia, Rome, Italy.

*Corresponding author. E-mail address: alberto.vaiarelli@gmail.com (A. Vaiarelli). https://doi.org/10.1016/j.rbmo.2024.103935 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS

Academic research evaluation
Journalist versus scientist
Quality versus quantity of publications
Quantitative metrics
Researchers

achieve individual goals in reaching a gradual in-depth knowledge and experience. The emphasis on accumulating citations becomes more critical than developing and testing novel ideas, and metrics may not adequately account for indicators of scientific soundness.

The temptation to fragment research findings into multiple publications or to favour systematic reviews and meta-analyses sacrificing quality for quantity becomes palpable. The consequence is a body of work that, while voluminous, lacks the depth and originality necessary for genuine academic impact (*Ioannidis et al., 2023*).

The overreliance on quantitative metrics raises questions about the effectiveness of evaluation systems. Relying heavily on medians and indicators such as citation counts may direct young researchers towards an exclusive focus on the number of published manuscripts, potentially neglecting the improvement of techniques or the sharing of scientific results beneficial to the community.

Moreover, these metrics may unintentionally overlook essential dimensions of an academic profile, such as real clinical activity, teaching-oriented work with patients, knowledge dissemination through scientific societies, practical and theoretical experiences in various settings, and the predisposition of young researchers to share their knowledge. In evaluating researchers, it is desirable that indicators go beyond numerical quantification and reflect the researcher's ability to fulfil the multifaceted actions characterizing an academic profile: research, teaching and mentoring.

Addressing these issues requires a re-evaluation of academic assessment criteria. Institutions should recognize the importance of quality over quantity and adopt a holistic approach that considers the depth, meaning and innovation in every scientific work. This approach could cultivate an environment valuing the diverse skills needed for excellence in academia.

In conclusion, while quantitative metrics have simplified aspects of young researcher evaluation, they have simultaneously given rise to a counterproductive culture where the volume of publications prevails over their substantial quality. Striking a balance between quantity and quality is crucial to fostering an academic environment that encourages meaningful contributions, innovation and the advancement of knowledge.

DATA AVAILABILITY

No data was used for the research described in the article.

REFERENCES

Conroy, G., 2024. Number of 'extremely productive' authors concerns scientists. Nature 625. https://doi.org/10.1038/d41586-023-03865-y.

Ioannidis, J.P.A., Collins, T.A., Baas, J., 2023. Evolving patterns of extremely productive publishing behavior across science. bioRxiv. https://doi.org/10.1101/2023.11.23.568476.

Ioannidis, J.P.A., Klavans, R., Boyack, K.w., 2018. The scientists who publish a paper every five days. Nature 561. https://doi.org/10.1038/d41586-018-06185-8.

Received 10 January 2024; accepted 19 February 2024.

RBMO

COMMENTARY

High-impact journal publishing: the devil is in the detail!

Barbara Lawrenz[1,2,*], Peter Humaidan[3], Christophe Blockeel[4,5], Juan-Antonio Garcia-Velasco[6,7], Human M. Fatemi[1]

ABSTRACT

Research in medicine is an indispensable tool to advance knowledge and improve patient care. This may be particularly true in the field of human reproduction as it is a relatively new field and treatment options are rapidly evolving. This is of particular importance in an emerging field like 'human reproduction', where treatment options evolve fast.The cornerstone of evidence-based knowledge, leading to evidence-based treatment decisions, is randomized controlled trials as they explore the benefits of new treatment approaches. The study design and performance are crucial and, if they are carried out correctly, solid conclusions can be drawn and be implemented in daily clinical routines. The dissemination of new findings throughout the scientific community occurs in the form of publications in scientific journals, and the importance of the journal is reflected in part by the impact factor. The peer review process before publication is fundamental in preventing flaws in the study design. Thus, readers of journals with a high impact factor usually rely on a thorough peer review process and therefore might not question the published data. However, even papers published in high-impact journals might not be free of flaws, so the aim of this paper is to encourage readers to be aware of this fact and critically read scientific papers as 'the devil lies in the details'.

Research is the cornerstone of knowledge acquisition and builds the foundation of advancements and improvements of treatment options for our patients. Publications in high-impact journals provide readers with the assurance that a study is ethically sound, skilfully planned and carried out. To assess the quality of a study and its findings, manuscripts undergo a review process by peer experts. In this process, the novelty and importance of the study, validity of the methods, including statistical analysis, quality of writing, data presentation and possible correlation between the study findings and the existing literature is scrutinized.

An indispensable tool of research, often called the 'gold standard', is randomized controlled trials (RCT). Their widely accepted experimental designs with the random assignment of participants into study and control groups aim to minimize selection bias, ensuring that groups are comparable once the study is finalized. This design should make it possible to draw causal inferences about the effects of a specific intervention by controlling for potential confounding variables. A further important aspect in the planning and performance of studies is 'equipoise'. When designing an RCT, clinical equipoise refers to the presumption that non-intervention (for neither the experimental group nor the control group) is the 'better one'. When there is no solid foundation on which to choose between two or more possibilities for treatment, a true state of equipoise exists.

A manuscript will undergo peer review before it is given consideration for publication and, if it is deemed to be of adequate quality, the manuscript will be accepted for publication, usually after revisions have been performed to address reviewers' comments. The objective of this procedure is to meticulously assess the paper's content for accuracy, quality and importance before disseminating it to the scientific community. Even though this process is contemporarily considered the gold standard for the evaluation of new scientific knowledge, it has several weaknesses (*Manchikanti et al., 2015*), among them lack of standardization of the peer review and lack of training in peer reviewing for the reviewers (*Blockeel et al., 2017*). Moreover, the number of peers with scientific training who are not only ready to devote their time to this labour-intensive procedure, but also competent to conduct a careful review might not meet the demands due

[1] IVF Department, ART Fertility Clinic, Abu Dhabi, UAE.
[2] Department of Reproductive Medicine, UZ Ghent, Ghent, Belgium.
[3] The Fertility Clinic, Skive Regional Hospital, Faculty of Health, Aarhus University, Skive, Denmark.
[4] Brussels IVF, Universitair Ziekenhuis Brussel, Brussels, Belgium.
[5] Department of Obstetrics and Gynaecology, University of Zagreb School of Medicine, Zagreb, Croatia.
[6] IVIRMA Global Research Alliance, IVIRMA Madrid, Madrid, Spain.
[7] IVIRMA Global Research Alliance, IVI Foundation, Instituto de Investigación Sanitaria La Fe, Valencia, and Rey Juan Carlos University, Madrid, Spain.

KEY WORDS
Critical assessment
High-impact journals
Randomized controlled trials
Research

*Corresponding author. E-mail address: barbara.lawrenz@artfertilityclinics.com (L. Barbara). https://doi.org/10.1016/j.rbmo.2024.103936 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

to the huge increase in scientific publications during the last few decades (*Ioannidis et al., 2023*).

The abovementioned points represent the fundamentals of research and the peer review process on which readers rely when reading a publication to acquire new knowledge and adjust their clinical handling. The purpose of the current paper is to illustrate potential weaknesses in the peer review process and highlight how the critical reading of papers remains important even when they are published in 'high-impact journals'.

To do this, we will consider a publication by Zaat and colleagues (*Zaat et al., 2023*), which has previously been commented on by Mackens and Blockeel (*Mackens and Blockeel, 2023*) and published in *The Lancet*. In an open-labelled, nationwide, multicentre, non-inferiority RCT, two different approaches to endometrial preparation for frozen embryo transfer (FET) were compared. This is a timely and important topic, especially as Zaat and colleagues (*Zaat et al., 2023*) explore a way to simplify the natural cycle endometrial preparation approach, which has recently been suggested to be superior to the medicated approach due to lower obstetric risks. The authors compared the ongoing pregnancy rates, achieved by the home-based and less costly detection of ovulation through the use of urinary LH kits versus a hospital-based approach, in which ovulation was triggered using recombinant human chorionic gonadotrophin. Based on their findings, the authors concluded that home-based ovulation monitoring is non-inferior to the hospital-controlled monitoring of ovulation for scheduling embryo transfer.

In a natural cycle approach, the detection of ovulation is crucial for planning an embryo transfer procedure as the endometrium is receptive approximately 120 h after progesterone secretion by the corpus luteum. Under the influence of progesterone, the endometrium undergoes secretory and functional transformation, which signifies its transition from the proliferative phase to the secretory phase. A surge of LH commonly precedes the progesterone rise; however, the time interval between the LH surge and the increase in progesterone can vary widely, and the LH peak can even coincide with the progesterone increase (*Coughlan et al., 2023*; *Erden et al., 2022*). As a consequence of the lack of any serum-based hormonal testing in the group of home-based patients, the timing of the progesterone rise remained unclear in the publication.

Another notable aspect of Zaat and colleagues' paper (*Zaat et al., 2023*) is that the study design does not appear to accommodate the possibility of premature ovulation in the hospital-monitored approach. Without first measuring LH, ovulation was induced at a predetermined follicle size between 16 and 20 mm. This implies that ovulation and the LH surge may have already started and gone undetected. Urinary LH testing has shortcomings regarding the correct identification of an LH surge due to either a delayed increase in the urine due to urinary clearance, a failure to detect the LH surge or an LH surge lacking ovulation and a progesterone rise; however, regular urinary LH testing in the home-based approach would probably be able to detect the start of ovulation/premature ovulation and would allow embryo transfer planning according to the cycle's specific characteristic. Importantly, undetected premature ovulation in the hospital-monitored group might have resulted in an asynchrony between the endometrium and the embryonic developmental stage, and might thus ultimately have led to a reduced pregnancy rate in this group.

Further points of criticism in terms of the different types of embryo cryopreservation and missing data on embryo quality are described, in addition to the main points detailed here, in the comment by Mackens and Blockeel (*Mackens and Blockeel, 2023*).

The current Commentary aims not only to highlight how shortcomings might be evident in studies published in high-impact journals, but also to illustrate how critical reading can reveal high quality. For instance, in a recent RCT (*Doyle et al., 2022*), the authors explored the effect on live birth rates of the timing of embryo transfer, using endometrial receptivity testing versus a standard timing of FET. The study showed that the timing of FET using endometrial receptivity testing did not increase live birth rates in single euploid blastocyst transfer. From their results, the authors were able to rightfully conclude that endometrial receptivity testing should not be recommended for routine use. The study is commendable for its approach to addressing a topic that is still up for debate. By following a consistent methodology before randomization, the study was able to establish group comparability and prevent bias since the intervention under investigation was the sole difference that distinguished the study group from the control group, which allowed clear conclusions to be drawn.

In summary, a constant pursuit of the knowledge inherent in science is crucial, as it remains the most reliable and progressive path forward for improving treatment options for our patients. The review process is implemented to scrutinize new findings; however, as this process is not perfect, readers of scientific papers — even in high-impact journals — need to be aware and critical of the limitations and potential biases inherent in scientific publications, as the devil definitely lies in the detail.

DATA AVAILABILITY

No data was used for the research described in the article.

REFERENCES

Blockeel, C., Drakopoulos, P., Polyzos, N.P., Tournaye, H., García-Velasco, J.A., 2017. Review the 'peer review'. Reprod. Biomed. Online 35, 747–749. https://doi.org/10.1016/j.rbmo.2017.08.017.

Coughlan, C., Ata, B., Gallego, R.D., Lawrenz, B., Melado, L., Samir, S., Fatemi, H., 2023. Interindividual variation of progesterone elevation post LH rize: implications for natural cycle frozen embryo transfers in the individualized medicine era. Reprod. Biol. Endocrinol. RBE 21, 47. https://doi.org/10.1186/s12958-023-01096-4.

Doyle, N., Jahandideh, S., Hill, M.J., Widra, E.A., Levy, M., Devine, K., 2022. Effect of Timing by Endometrial Receptivity Testing vs Standard Timing of Frozen Embryo Transfer on Live Birth in Patients Undergoing In Vitro Fertilization: A Randomized Clinical Trial. JAMA 328, 2117–2125. https://doi.org/10.1001/jama.2022.20438.

Erden, M., Mumusoglu, S., Polat, M., Yarali Ozbek, I., Esteves, S.C., Humaidan, P., Yarali, H., 2022. The LH surge and ovulation re-visited: a systematic review and meta-analysis and implications for true natural cycle frozen thawed embryo transfer. Hum. Reprod. Update. https://doi.org/10.1093/humupd/dmac012 dmac012.

Ioannidis, J.P.A., Pezzullo, A.M., Boccia, S., 2023. The Rapid Growth of Mega-Journals: Threats and Opportunities. JAMA 329, 1253. https://doi.org/10.1001/jama.2023.3212.

Mackens, S., Blockeel, C., 2023. Home-based monitoring prior to frozen embryo transfer: the new gold standard? The Lancet 402, 1304–1306. https://doi.org/10.1016/S0140-6736(23)01798-1.

Manchikanti, L., Kaye, A.D., Boswell, M.V., Hirsch, J.A., 2015. Medical journal peer review: process and bias. Pain Physician 18, E1–E14.

Zaat, T., de Bruin, J.-P., Goddijn, M., van Baal, M., Benneheij, S., Brandes, M., Broekmans, F., Cantineau, A., Cohlen, B., van Disseldorp, J., Gielen, S., Groenewoud, E., van Heusden, A., Kaaijk, E., Koks, C., de Koning, C., Klijn, N., van der Linden, P., Manger, P., Moolenaar, L., van Oppenraaij, R., Pieterse, Q., Smeenk, J., Visser, J., van Wely, M., Mol, F., 2023. Home-based monitoring of ovulation to time frozen embryo transfers in the Netherlands (Antarctica-2): an open-label, nationwide, randomized, non-inferiority trial. Lancet Lond. Engl. 402, 1347–1355. https://doi.org/10.1016/S0140-6736(23)01312-0.

Received 1 February 2024; received in revised form 17 February 2024; accepted 19 February 2024.

RBMO

COMMENTARY

Mother, father, son and the Italian Law 40/2004. No 'delete' key

Emanuela Turillazzi[a], Donato Morena[b], Luigi Papi[a], Vittorio Fineschi[b,*]

ABSTRACT

Assisted reproductive technology (ART) has emerged in recent years as a point of significant innovation in the medical field but is also controversial from a bioethical and legal standpoint. In the Italian context, this matter is regulated by Law 40/2004, which specifically requires that informed consent should be obtained from both members of a couple before proceeding with any ART procedure. This consent is deemed irrevocable at the moment of egg fertilization. Recently, a ruling by the Italian Constitutional Court on this matter elicited controversy. The decision permitted embryo transfer even in a case of parental separation, notwithstanding the father's explicit opposition. The Court emphasized the priority of the woman's psychophysical health over the man's, highlighting the traumatic consequences of interrupting the undertaken path. As a result, both the man's right to self-determination regarding the decision to become a father and the need for informed consent at every stage of medical procedures have been downplayed. Moreover, the extensive utilization of procedures like embryo cryopreservation, with associated parental implications, particularly concerning the time frame and the actuality of informed consent, is posing challenges to the initial application framework of Law 40/2004. The objective of this Commentary is to scrutinize and discuss the issues mentioned above.

INTRODUCTION

Assisted reproductive technology (ART) predominantly revolves around continuous advancements in the biomedical context and the associated ethical, social and legislative challenges (*Cobb and Ke, 2018*). The critical intersections between ethics and law are notably concentrated on: (i) acceptable family building, encompassing considerations such as post-mortem insemination and lesbian couples; (ii) ensuring patient safety through the regulation of centres and adherence to quality standards; (iii) addressing issues of justice and equality, particularly in the context of treatment reimbursement; and (iv) safeguarding the welfare of the child, involving aspects such as donor anonymity, health considerations and embryo transfer (*Pennings, 2009*). Furthermore, various factors influence the diverse regulatory approaches adopted by different countries.

Within this complex scenario, the association between ART and informed consent becomes a potential challenge. In the Italian context, this point is governed by Law 40/2004. A contentious issue arises with the provision in Italian law, stipulating that consent becomes irrevocable upon fertilization of the egg. As the Italian Constitutional Court has recently adjudicated on this matter, this paper aims to critically examine and discuss the implications of the Court's ruling.

INFORMED CONSENT, ITALIAN LAW AND COURTS

Since its enactment, Law 40/2004 has undergone multiple amendments. Over the years, prohibitions on producing a maximum number of embryos during the same procedure and conducting pre-implantation diagnosis have been lifted, permitting the cryopreservation of embryos. Moreover, the Italian Constitutional Court scrutinized the prohibition of research activities concerning embryos and the irrevocability of consent, asserting that the 'tragic choice' between respecting life at its inception and scientific research, which poses ethical and juridical divisions, should be determined by Parliament (*Italian Constitutional Court, July 2023*, n.161).

Law 40/2004 highlights the importance of informed consent (Article 6), mandating mutual agreement from both members of a couple before proceeding with any ART procedure. The doctor overseeing the procedure at the centre is obliged to provide information to the subjects on both the medical and legal aspects of the methodologies. This agreement must be documented in writing, with the signature of the responsible doctor at the respective centre.

A waiting period of no less than 7 days is mandated between the expression of the desire to undergo ART and the actual application of the technique. During this

[a] Department of Legal Medicine, University of Pisa, Pisa, Italy.
[b] Department of Anatomical, Histological, Forensic and Orthopedic Science, Sapienza University of Rome, Rome, Italy.

*Corresponding author. E-mail address: vittorio.fineschi@uniroma1.it (V. Fineschi). https://doi.org/10.1016/j.rbmo.2024.103970 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS
Assisted reproduction
Embryo transfer
Father
Informed consent
Italian law

period, either party has the right to revoke their wish until the insemination of the oocyte takes place. A decree issued by the Italian Ministry of Justice on 28 December 2016 (*Italian Ministry of Justice, 2016*) reiterated key points of consensus in ART procedures, including the stipulation that applicants can only revoke their consent up to the moment of fertilization of the eggs.

The matter of the irrevocability of consent provided by the couple was addressed in 2021 by a Southern Italian court (*Italian Court of Santa Maria Capua Vetere, 2021*), which granted authorization for the transfer of an embryo into the maternal womb. This decision was made despite the occurrence of separation and the explicit objection of the male partner to the embryo transfer. In that case, the judge's ruling rested on two primary considerations. First, it prioritized the protection of the cryopreserved embryo, highlighting its inherent inclination towards life and developmental potential. As a result, the decision endorsed the embryo transfer. Second, the judgment took into account the principle of self-responsibility and the legitimate expectation expressed in the consent given by both parents. The breakdown of the family and marital relationship, which originally justified the parental project the couple intended to undertake, was deemed irrelevant by the judges (*Di Fazio et al., 2021*). Following this decision, there were further jurisprudential developments, ultimately culminating in a 2023 judgment rendered by the Constitutional Court (*Italian Constitutional Court, July 2023*, n.161).

In summary, a couple mutually decided to undergo ART procedures at an Italian health facility. Upon verification of the requirements, informed consent was obtained from both members of the couple. Subsequently, their gametes were collected, and viable embryos suitable for transfer were produced. However, due to suboptimal endometrial receptivity, the procedure was temporarily suspended, and the embryos were cryopreserved.

Subsequently, the couple became entangled in a highly contentious separation. In this context, the woman opted to proceed with the ART procedure, requesting the medical centre to proceed with the transfer of the cryopreserved embryos. In the interim, the husband, following the formal initiation of the judicial separation process, officially withdrew his consent to the ART procedure. As a result, the medical centre rejected the woman's request and she appealed to the Ordinary Court of Rome. The argument of her appeal emphasized that Article 6 paragraph 3 of Law No. 40/2004 prohibits the revocation of consent in the period following fertilization. Additionally, it asserted that the 'right' to become a mother is an absolute and fundamental right protected by the Italian Constitution under Articles 2, 31 paragraph 2, and 32.

The Ordinary Court of Rome, referring to Articles 2, 3, 13 paragraph 1 and 32 paragraph 2, and 117 paragraph 1 of the Constitution, the latter in relation to Article 8 of the European Convention on Human Rights, posed questions regarding the constitutional legitimacy of Article 6, paragraph 3, last decision, of Law 40/2004. According to the referring judge, the contested provision could potentially infringe upon the right to choose regarding the assumption of the parental role, particularly in cases where the request for implantation is made under 'a different legal situation' from that existing at the time of expressing the will.

The judge argued that the provision might prejudice the right to choose if, considering the passage of time, the implantation is requested under circumstances different from those originally contemplated. Therefore, since Article 5, paragraph 1, of Law No. 40/2004 allows access to ART 'only to adult couples of different sexes, married or cohabiting, of potentially fertile age, both living', if 'the couple's plan ceases before the transfer of the embryo' the 'revocation of consent' should be considered permissible.

Thereafter, the Constitutional Court ruled that the questions regarding the constitutional legitimacy of the contested Article 6, paragraph 3, last decision, of Law 40/2004 were unfounded.

KEY PASSAGES FROM THE RULING

With its decision, the Italian Constitutional Court consolidated the legal provision establishing the irrevocability of consent expressed by a couple undergoing an ART procedure. The Court gives precedence to the psycho-physical health of the woman, underscoring the traumatic consequences that the interruption of the pursued path would inflict on her, particularly when the process has advanced to the fertilization stage. According to the Court, this impact becomes particularly pronounced:

if the woman's age or physical conditions, coupled with the elapsed time since the cryopreservation of the contested embryo, hinder her ability to initiate a new ART path, with an absolute preclusion, at this point, of her freedom of self-determination regarding procreation (*Italian Constitutional Court, July 2023*, n.161).

In the words of the Court:

In accessing ART, the woman is immediately engaged with her own body, a factor incomparably more significant than what occurs for the man. Indeed, in order to achieve the shared parenting project, she is first subjected to demanding cycles of ovarian stimulation, during which the onset of pathologies, including serious ones, cannot be excluded (*Italian Constitutional Court, July 2023*, n.161).

Furthermore, access to ART imposes:

a significant burden on women, requiring them to make their own corporeality available with a substantial physical and emotional investment in the context of parenthood, involving risks, expectations, and suffering, and which has a turning point when one or more embryos are formed. The body and mind of the woman are therefore inseparably involved in this process, which culminates in the concrete hope of conceiving a child through the implantation of the embryo in her own uterus. (*Italian Constitutional Court, July 2023*, n.161).

It is worth noting that, although the moment of embryo implantation evidently exclusively affects the woman's body, the subsequent phases of pregnancy, childbirth and fatherhood (as a consequence of that implantation) also impact the psycho-physical health of the man. Parenthood, particularly for first-time fathers, represents an unfamiliar and unpredictable event beyond their usual sphere of control. This experience can unveil numerous potential stressors, such as evolving relationships with the mother and additional financial responsibilities and concerns, which may impact the capacity to be a competent parent (*Kotelchuck, 2022*). Moreover, parenting after pregnancies resulting from ART can be challenging and emotionally demanding

even for fathers (*El Kissi et al., 2013*), and it could be more severe if pregnancies are involuntary and no longer desired.

Another right under consideration is the man's right to self-determination regarding the decision of whether or not to become a father. The Constitutional Court emphasizes the importance of the context in which he becomes aware of the potential use of cryopreservation and provides consent to it. This involves a specific assumption of responsibility regarding parentage, leading to the attribution of filiation status to the offspring, regardless of subsequent events in the couple's relationship.

The decision suggests a form of self-determination by the father that is not unconditional but rather functional and directed towards assuming parental responsibility, from which he cannot escape once fertilization has occurred. In essence, it reflects a consent whose consequences unfold over a significant period, compelling the man to 'become the father of a child, even when the conditions under which he initially shared the parental project no longer exist' (*Italian Constitutional Court, July 2023*, n.161).

In terms of informed consent, in 2009 the Italian Constitutional Court repealed the prohibition on embryo cryopreservation that was initially outlined in the original draft of the law (*Italian Constitutional Court, 2009*). Following the removal of this ban, there was an 'expansion of the cryopreservation technique with the consequent possibility of a temporal split between fertilization and implantation', as mentioned by the Court in a passage of the discussed judgment (*Italian Ministry of Health, 2023*).

This statement encapsulates another issue in the informed consent process, emphasizing the profound changes that have transpired over the years in ART procedures. The recent surge in frozen-thawed embryo replacement (FER) cycles and advancements in embryo cryopreservation techniques have led to outcomes for FER that are comparable to fresh embryo transfer (*Noble and Child, 2020*). Additionally, the elective freezing of all high-quality embryos with subsequent transfer in later cycles has witnessed a notable increase in recent times (*Roque et al., 2019*).

Regarding Italy, the scenario outlined in the latest Report from the Minister of Health to Parliament on the Application of Law 40/2004 emphasizes that FER cycles have grown in both absolute terms, going from 508 cycles in 2008 to 27,204 cycles in 2021, and percentage terms, increasing from 1.1% of all techniques in 2008 to 34.5% in 2021. The percentage of application of fresh techniques, however, has decreased overall, from 89.1% in 2005 to 63.7% in 2021 (*Italian Ministry of Health, 2023*).

The increasing and tangible possibility of embryo transfer occurring at a temporal distance from the initiation of the procreative process raises significant considerations. The rising prevalence of cryopreserved embryos worldwide implies that the window for decision making could be substantially extended, spanning from days to years. Notably, it underscores the concerns expressed by the presiding judge about 'the fact that the legislator did not attach importance to the events following fertilization'. The progressive and rapidly evolving advancements in embryo cryopreservation have significantly altered the original application landscape of Law 40/2004 (*Italian Constitutional Court, July 2023*, n.161).

Finally, the entire matter is further complicated by the completion of the procreative process without the man's consent, especially in the face of the explicit revocation of consent. This not only contradicts the protection of the man's right to health and his freedom of self-determination to the extent that the filiation project is framed as coercive, but also conflicts with the fundamental principles regarding informed consent within the Italian legislative context. Italian law 219/2017 ('Provisions for informed consent and advance treatment directives') stipulates in Article 1 that no medical treatment can be initiated or continued without the free and informed consent of the individual, except in cases expressly provided for by the law.

Similar to any other medical treatment, assisted reproduction necessitates the disclosure of the intervention's purpose, the intended procedures, associated risks, benefits and available alternatives. Specific issues that have to be discussed in this field regard the legal implications for the woman, the man and the unborn child. Comprehensive disclosures and discussions within the informed consent process are imperative at every phase of implementing medically assisted procreation techniques.

A very thorny point is the provision of Italian law stipulating that consent becomes irrevocable once the fertilization of the egg occurs. Informed consent should be obtained at each stage of any healthcare treatment, and this principle ought to be consistently applied throughout every phase of the ART procedure for both parents. This becomes particularly relevant given the potential temporal gap between the initiation of the procedure and embryo transfer, resulting from the increasing utilization of embryo cryopreservation strategies (*Cohen and Adashi, 2016*).

In the delicate and aporic situation at hand, it is not feasible to reconcile all the conflicting interests involved in the case: (i) the safeguarding of the woman's psycho-physical health and her right to pursue motherhood; (ii) the man's freedom of self-determination and his will not to become a father; and (iii) for the embryo, the right to life and family (*Italian Constitutional Court, July 2023*, n.161).

In conclusion, the Italian Constitutional Court upholds the rights of the woman and those of the embryo over those of the man, asserting that such prioritization does not constitute a violation of the principle of equality. The ruling is based on the interpretative premise that a woman retains the right to refuse the transfer of the embryo to her uterus, as the rule cannot endorse coercive embryo transfer to woman's womb. We find this compression of paternal self-determination too draconian and unconvincing.

DATA AVAILABILITY

No data was used for the research described in the article.

REFERENCES

Cobb, L.N., Ke, R.W., 2018. Ethical considerations in the field of assisted reproductive technology. Minerva endocrinologica 43 (1), 80–86. https://doi.org/10.23736/S0391-1977.17.02664-5.

Cohen, IG, Adashi, EY., 2016. Embryo Disposition Disputes: Controversies and Case Law. The Hastings Center Report 46 (4), 13–19. https://doi.org/10.1002/hast.600.

Di Fazio, N., Fineschi, B., Caporale, M., Del Fante, Z., Volonnino, G., Santoro, P., La Russa, R., 2021. Recent Judgement of the Italian Judiciary about medical assisted procreation (MAP): is informed consent valid after parents separation? La Clinica Terapeutica 172 (4), 253–255. https://doi.org/10.7417/CT.2021.2325.

El Kissi, Y., Romdhane, A.B., Hidar, S., Bannour, S., Ayoubi Idrissi, K., Khairi, H., Ben Hadj Ali, B., 2013. General psychopathology, anxiety, depression and self-esteem in couples undergoing infertility treatment: a comparative study between men and women. European journal of obstetrics, gynecology, and reproductive biology 167 (2), 185–189. https://doi.org/10.1016/j.ejogrb.2012.12.014.

Kotelchuck, M., 2022. The impact of fatherhood on men's health and development. Engaged Fatherhood for Men, Families and Gender Equality: Healthcare, Social Policy, and Work Perspectives 63–69.

Italian Constitutional Court, May 2009, n.151

Italian Constitutional Court, July 2023, n.161

Italian Ministry of Health. (2023). Report of the Minister of Health to Parliament on the Implementation Status of the Law Containing Rules on Medically Assisted Procreation (Law of February 19, 2004, No. 40, Article 15) - Activities of Assisted Reproductive Technology Centers Year 2021. Last access on 07 February 2024 at: https://www.salute.gov.it/imgs/C_17_pubblicazioni_3380_allegato.pdf

Italian Court of Santa Maria Capua Vetere, First Civil Section, Order of January 27, 2021. Available at: https://www.biodiritto.org/Biolaw-pedia/Giurisprudenza/Tribunale-di-Santa-Maria-Capua-Vetere-prima-sezione-civile-ordinanza-27-gennaio-2021-e-possibile-usare-gli-embrioni-crioconservati-contro-la-volonta-dell-ex-partner (last access on 03 May, 2024).

Italian Ministry of Justice (2016). Regulation containing rules on the expression of the will to access medically assisted procreation techniques, in implementation of Article 6, paragraph 3, of Law February 19, 2004, No. 40. Last access on 07 February 2024 at: https://www.iss.it/documents/20126/0/DECRETO+28+dicembre+2016+%2C+n.+265+..pdf/e515a828-2ba3-67de-0620-947a291ca8c5?t=1610640451851

Noble, M., Child, T., 2020. The role of frozen–thawed embryo replacement cycles in assisted conception. The Obstetrician & Gynaecologist 22 (1), 57–68.

Pennings, G., 2009. International Evolution of Legislation and Guidelines in Medically Assisted Reproduction. Reproductive BioMedicine Online 18, 15–18. https://doi.org/10.1016/s1472-6483(10)60443-9.

Roque, M., Haahr, T., Geber, S., Esteves, S.C., Humaidan, P., 2019. Fresh versus elective frozen embryo transfer in IVF/ICSI cycles: a systematic review and meta-analysis of reproductive outcomes. Human reproduction update 25 (1), 2–14. https://doi.org/10.1093/humupd/dmy033.

Received 10 February 2024; received in revised form 11 March 2024; accepted 18 March 2024.

RBMO

ELSEVIER

ARTICLE

Anticentromere antibodies are the most potent antinuclear antibodies in reducing live birth outcomes after ICSI

BIOGRAPHY

Dr. Shokichi teramoto is Director of the Tawako Dream Clinic Group. He found that oocytes retrieved from small follicles are often mature and suitable for IVF, and that there is a substantial possibility of retrieving oocytes from prematurely ruptured follicles.

Shokichi Teramoto[a], Tsuyoshi Ueno[b], Fumihito Aono[a], Tsuyoshi Okubo[b], Tomoya Segawa[a,b], Hisao Osada[a], Makio Shozu[a,c,*]

KEY MESSAGE

Anticentromere antibodies (ACA) are the most potent of all antinuclear antibodies in reducing ICSI outcomes. Fertilization with multiple pronuclei formation and subsequent cleavage failure is pathognomonic for infertility related to ACA. This condition may be recognized as a distinct clinical entity of refractory infertility caused by ACA.

ABSTRACT

Research question: How, and to what extent, do anticentromere antibodies (ACA) reduce live birth outcomes after ICSI?

Study design: Retrospective cohort study of infertile women aged 30–43 years who underwent ICSI between September 2016 and March 2021. Women with a history or current diagnosis of symptomatic connective tissue disease were excluded. Immunofluorescence staining detected antinuclear antibodies (ANA). Staining pattern and titre (cut-off, 1:160) were used to divide infertile women into three groups: positive for ACA (ACA+) ($n = 28$); positive for ANA other than ACA (ANA+) ($n = 77$); and negative for both ACA and ANA (control) ($n = 3723$).

Results: Cumulative live birth rate (CLB) was lowest in ACA+ (7%, 31% and 46% in ACA+, ANA+ and control, respectively) (ACA+ versus control, $P < 0.0001$; ACA+ versus ANA+, $P = 0.011$; ANA+ versus control, $P = 0.012$). A small impairment in meiosis I and a larger impairment in meiosis II, fertilization and embryo cleavage caused the decrease. Multiple pronuclei formation increased (RR versus control, 5.33; 95% CI 4.26 to 6.65) and good-quality blastocyst development decreased (RR 0.34; 95% CI 0.22 to 0.53). Multiple logistic regression analysis showed that ACA was associated with CLB outcome (OR 0.08, 95% CI 0.02 to 0.36); the other four ANA staining patterns were not.

Conclusions: The effect of ACA on live birth outcomes is strongest after ICSI among ANA, primarily through the impairment of meiosis II and subsequent stages. Repeated ICSI failure and eggs with multiple pronuclei may warrant specific testing for ACA.

[a] Natural ART Clinic at Nihonbashi, 2-7-1, Nihonbashi, Chuo-ku, Tokyo, Japan, 103-6008.
[b] Shimbashi Yume Clinic, 2-5-1, Shimbashi, Minato-ku, Tokyo, Japan, 105-0004.
[c] Evolution and Reproduction Biology, Medical Mycology Research Center, Chiba University, 1-8-1, Inohana, Chuo-ku, Chiba, Chiba Prefecture, Japan, 260-8673.

*Corresponding author. E-mail address: shozumakio@chiba-u.jp (M. Shozu). https://doi.org/10.1016/j.rbmo.2024.103864

Declaration The authors report no financial or commercial conflicts of interest.

KEYWORDS
anticentromere antibody
antinuclear antibody
intracytoplasmic sperm injection
multiple pronucleus

INTRODUCTION

Antinuclear antibodies (ANA) are implicated in the pathogenesis of many systemic autoimmune diseases, such as systemic lupus erythematosus, which also affects fertility (*Simopoulou et al., 2019*; *Zeng et al., 2019*; *Cavalcante et al., 2020*). The impairment of fertility has also been observed in IVF and embryo transfer outcomes (*Kikuchi et al., 2003*; *Ying et al., 2012*; *2013a*; *Zhu et al., 2013*; *Fan et al., 2017*; *Ticconi et al., 2023*), suggesting that even recent assisted reproductive technology is unable to overcome this condition. Antinuclear antibodies consists of a large group of autoantibodies targeting a variety of nuclear antigens, which may play a role in human reproduction. Evidence is limited on which type of ANA is responsible for infertility (*Li et al., 2015*).

Nuclear antibodies have been studied to some extent. *Li et al. (2015)* examined IVF, embryo transfer and intracytoplasmic sperm injection (ICSI) outcomes, and reported associations between anti-Scl-70 and anti-PM-Scl antibodies with decreased pregnancy ratio, and associations between anti-Jo-1, anti-Rib-p and anti-double strand DNA (anti-dsDNA) antibodies with early miscarriage (*Li et al., 2015*). *Fan et al. (2017)* found anti-dsDNA antibodies to be associated with a decreased number of retrieved oocytes and high-quality embryos in the case of IVF. Furthermore, the implantation rate and clinical pregnancy ratios decreased, and the miscarriage ratio increased in women who were positive for anti-dsDNA antibodies (*Fan et al., 2017*). Compared with the patient group positive for ANA other than anti-dsDNA antibodies, anti-dsDNA antibodies seem to have the most significant negative effect on IVF outcomes; however, no studies have definitively confirmed this.

The anticentromere antibody (ACA) is another ANA that may be associated with reduced IVF, embryo transfer and ICSI outcomes. *Shirota et al. (2011)* compared ICSI outcomes between eight women with ACA and 58 women with ANA other than ACA. The maturation of retrieved oocytes and the cleavage of fertilized eggs decreased by 20% for each, whereas fertilization was unaffected. Similarly, *Ying et al. (2013b)* evaluated IVF outcomes in 20 women with ACA, 51 women with ANA other than ACA and 116 age-matched women without ACA and ANA. The ACA was associated with a 20% reduction in oocyte maturation and cleavage of fertilized eggs, whereas the fertilization rate was not reduced. In addition to embryo development, *Ying et al. (2013b)* also examined pregnancy outcome and found that implantation and clinical pregnancy ratios were reduced by about 50% in ACA-positive women compared with women without ANA. These results suggest that ACA is the most significant negative factor affecting pregnancy outcomes after ICSI by reducing oocyte maturation, embryo cleavage and implantation ratios, but not by reducing fertilization rates (*Ying et al., 2013b*). The evidence, however, is limited, and further research is needed to conclude that ACA is responsible for the most adverse effect on ICSI outcomes and to consider the underlying mechanisms for the impairment.

In the present study, the association between ACA and ICSI outcomes was investigated to determine whether ACA is the most significant predictor of the reduced outcomes of all antinuclear antibodies. Outcomes were also analysed by every developmental stage and the stage responsible for ACA's action were determined.

MATERIALS AND METHODS

Participants

The study population comprised infertile women who visited our private clinic and underwent ICSI between 1 September 2016 and 31 March 2021. All patients were offered a screening test for antinuclear autoantibodies before ICSI, and about 85% accepted the offer. Patients aged 30–43 years who underwent the screening test and one or more ICSI cycles were included in the analysis. Women with current disease or a history of connective tissue disease relevant to ANA, such as systemic lupus erythematosus or scleroderma, were excluded. A total of 3828 patients were included in the analysis (**Supplementary Figure 1**). Pregnancy outcomes after single embryo transfers were analysed as of March 2022.

Study design and specific aim

This retrospective cohort study investigated the relationship between ACA and live birth outcomes compared with other ANAs through relative risk and logistic regression analyses. Outcome reduction was assessed at every stage from oocyte retrieval until live birth to identify the critical stage for ACA-induced infertility.

Ethical approval

This study was approved by the Institutional Review Board (Towako2018-01a, 1 March 2019). Patients who consented to the study were included.

Fluorescence antinuclear antibody test

Serum antinuclear antibody was detected using the Fluoro HEPANA-2 test (MBL, Nagoya, Japan). The immunofluorescence staining pattern of Hep-2 cell nuclei was categorized according to the International Consensus on ANA Patterns: homogeneous, peripheral, speckled, nucleolar and discrete-speckled (*Chan et al., 2015*). Of these, the discrete-speckled pattern is unique to ACA and, therefore, is referred to as the centromere type, whereas other patterns correspond to several different antigens; therefore, the detected antigen cannot be specified. Staining intensity was quantified by titration. The cut-off titre used for a positive result in this study was 1:320. Patients were divided into three groups: ACA+ (ACA positive, $n = 28$), ANA+ (ACA negative and ANA positive, $n = 77$), and control (ACA negative and ANA negative, $n = 3723$). Two double-positive women (a woman with 1:320 ACA and 1:320 ANA and a woman with 1:1280 ACA and 1:640 ANA) were assigned to the ACA+.

Intracytoplasmic sperm injection

The handling and culture of oocytes and ICSI protocol were described previously (*Teramoto et al., 2016*; *2019*). Nasal buserelin acetate triggered LH surge (300 μg, twice in 1-h intervals). Oocytes were retrieved from both regular (≥10 mm) and small-sized follicles (4–9 mm) during mild stimulation cycles or natural cycles (*Teramoto et al., 2016*). Oocyte maturity was determined within 2 h of retrieval. If necessary, oocytes were matured *in vitro* in culture media for up to 24 h. Both genuine and MII oocytes that had been matured *in vitro* were fertilized by ICSI and cultured into blastocysts. The resulting good-quality blastocysts (≥3BB of Gardner's classification or ≥180 μm diameter on day 5 or 6) were frozen once and then each embryo was individually transferred one by one during natural or hormonally controlled cycles. Each oocyte was individually cultured in a single droplet (20 μl) of NAKA one-step medium® (NAKA Medical Co., Tokyo, Japan) in an atmosphere of 5% oxygen, 5% carbon

TABLE 1 PATIENT CHARACTERISTICS

Characteristics	Total	ACA+	ANA+	Control	*P*-value[a]
Women, *n*	3828	28	77	3723	
Age at initial ICSI, years (IQR)	38 (35–41)	38 (35.25–40.75)	37 (35.5–41)	38 (35–41)	0.970
AMH at initial ICSI, ng/ml	2.0 (0.9–3.7)	2.6 (1.7–3.9)	1.7 (1.0–3.3)	2 (0.9–3.7)	0.202
Total oocyte retrieval	12,000	148	243	11,609	
Oocyte retrieval per patient	2 (1–4)	4 (2–8.5)[b]	2 (1–4)[c]	2 (1–4)[c]	0.010
Retrieved oocyte per cycle	4 (2–8)	6 (3–9)[b]	3 (1–7)[c]	4 (2–8)c	<0.0001
Cause of infertility (% couples)					
Male factor	1237 (32)	10 (36)	20 (26)	1207 (32)	0.593
Female factor	362 (9)	3 (11)	8 (10)	351 (9)	0.444
Unknown causes	2320 (61)	16 (57)	49 (64)	2255	0.268
Oocyte stimulation cycles (% total oocyte retrieval)	3248 (27)	65 (43)	64 (26)	3121 (27)	<0.0001

Continuous values presented as median (interquartile range).

[a] Comparison among three groups.

[b,c] Subgroups with column ratios not significantly different at the 0.05 level after adjustment for multiple comparisons.

ACA, anticentromere antibody; ACA+, ACA positive; AMH, anti-Müllerian hormone; ANA, antinuclear antibody; ANA+, ANA positive; ICSI, intracytoplasmic sperm injection; IQR, interquartile range.

dioxide and 90% nitrogen. The Oocyte/Embryo Vitrification Kit and Cryotop® (Kitazato Medical Co., Shizuoka, Japan) were used for cryopreservation.

Outcomes measured

The primary outcome was cumulative live birth, indicating that the patient experienced a live birth at least once through the ICSI cycles carried out during the study period. Secondary outcomes included a cumulative ratio of one or more oocyte retrieved, embryo transfer, biochemical pregnancy (≥20 IU HCG on day 7 of embryo transfer), clinical pregnancy (gestational sac at 5 weeks and 1 day of pregnancy) and fetal heartbeat positivity (at 6 weeks and 4 days of pregnancy). Secondary outcome measures of oocyte development included the germinal vesicle breakdown, first polar body extrusion, two pronucei (2PN) formation, non-2PN formation and good-quality blastocyst. Both regular-sized and small-sized follicles were included in the analysis unless specified otherwise.

Statistics

JPM statistical software, version 16 (SAS Institute, Cary, NC, USA) was used for statistical analysis. The Wilcoxon rank sum and Kruskal–Wallis tests were used to detect differences among the three groups for continuous variables that did not show normal distribution. $P < 0.05$ was considered significant. For post-hoc tests, Bonferroni's method was used. The difference in proportion was detected by Fisher's exact probability test at the significance level of 0.05 and adjusted for multiple comparisons by Hochberg's method. Multiple logistic regression analysis was applied to correlate pregnancy outcomes with antinuclear staining patterns. The patterns were converted to binary data with the cut-off. Significant variables were selected using a forward stepwise method.

RESULTS

Patient characteristics

Of 3828 patients, 28 (0.73%) had an ACA titre above the cut-off (**Supplementary Figure 2**). Patient age, anti-Müllerian hormone levels and infertility causes were not different among the three groups; however, the ACA+ group had more treatment cycles (ACA+ versus ANA+, $P = 0.003$; ACA+ versus control, $P < 0.007$) and more ovarian stimulation cycles (ACA+ versus ANA+, $P = 0.001$; ACA+ versus control, $P < 0.0001$), which helped in achieving more oocytes (TABLE 1).

Pregnancy outcomes

The cumulative live birth rate was decreased in ACA+ compared with ANA+ ($P = 0.011$) and the control ($P < 0.0001$) (FIGURE 1).

To identify the stages responsible for the decline, cumulative outcome risks were

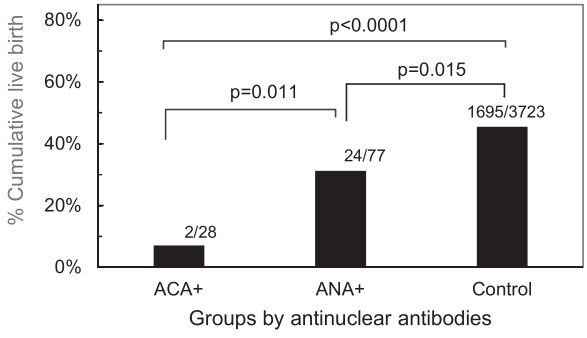


FIGURE 1 Cumulative live birth rate. Incidence of women who achieved at least one live birth during the treatment period after intracytoplasmic sperm injection was calculated on an intention-to-treat basis. The number on the right shoulder of each column represents the actual number of patients. *P*-values on the vertical line represents Fisher's exact probability between indicated columns; statistical differences are significant after adjustment for multiple comparison. ACA, anticentromere antibody; ACA+, ACA positive; ANA, antinuclear antibody; ANA+, ANA positive.

TABLE 2 PATIENT-BASED CUMULATIVE INTRACYTOPLASMIC SPERM INJECTION OUTCOMES EVALUATED AT EACH STAGE

Cumulative outcomes	ACA+				ANA+				Control				*P*-value[a]
	n	/	*n*	%	*n*	/	*n*	%	*n*	/	*n*	%	
One or more oocytes retrieved/number of patients	28	/	28	100	77	/	77	100	3705	/	3723	100	1.000
Blastocyst-transfer/≥1 oocyte retrieved	10	/	28	36[b]	51	/	77	66[c]	2684	/	3705	72[c]	0.000
Biochemical pregnancy/embryo-transfer	4	/	10	40[b]	39	/	51	76[b,c]	2314	/	2684	86[c]	0.000
Clinical pregnancy/biochemical pregnancy	4	/	4	100	36	/	39	92	2235	/	2314	97	0.265
Fetal heart beat positivity/clinical pregnancy	3	/	4	75	32	/	36	89	2084	/	2235	93	0.118
Live birth/feal heart beat positivity	2	/	3	67	24	/	32	75	1695	/	2084	81	0.320

[a] Comparison between three groups.

[b,c] A subgroup whose column ratio is not significantly different at the 0.05 level (Fisher's exact probability) after adjusting for multiple comparisons using Hochberg's method.

ACA, anticentromere antibody; ANA, antinuclear antibody.

analysed for every stage (TABLE 2). Oocyte retrieval was unaffected (all ACA+ patients had one or more oocytes), whereas blastocyst transfer and biochemical pregnancy were decreased compared with the control group, namely a decrease in blastocyte yield and implantation success risk. Statistical analysis on outcomes of clinical pregnancy and thereafter was inappropriate because of the small number of patients.

In the case of ANA+, no decrease in blastocyst yield or implantation success was observed compared with the control.

Embryo development

In-vitro development was evaluated for every step to determine the critical step for reduced blastocyst yield. In ACA+, first polar body extrusion (completion of meiosis I), fertilization with 2PN formation and embryo cleavage to good-quality blastocysts were decreased (TABLE 3). Relative risk to the control were 0.88 (95% CI 0.81 to 0.96), 0.57 (0.51 t 0.62) and 0.34 (0.22 to 0.53), respectively. Corresponding to the decrease in 2PN formation, non-2PN eggs increased so that the overall fertilization ratio remained unchanged (TABLE 3). Among non-2PN eggs, single pronucleus eggs and three pronuclei (3PN) eggs or higher increased (TABLE 4), although the increase was more pronounced in eggs with a higher number of pronuclei; the relative risk increased significantly with the number of pronuclei from 2.55 to 18.01 (**Supplementary Table 1**). The relative risk of 3PN eggs or higher was 5.33 (95% CI 4.26 to 6.65). Such increases were observed in ACA+, but not in ANA+.

In ANA+, a slight decrease in first polar body extrusion rate (RR versus control, 0.96) was observed, but otherwise no decrease at any step, unlike ACA+ (TABLE 3). Moreover, the number of pronuclei and their distribution were unchanged so that the relative risk remained close to 1 (TABLE 4 and **Supplementary Table 1**).

Antibody type-specific effect

Logistic regression analyses were conducted with cumulative live birth rate as the dependent variable and ANA staining patterns as independent variables.

TABLE 3 OOCYTE-BASED INTRACYTOPLASMIC SPERM INJECTION OUTCOMES

Developmental stages	Formula	ACA+		ANA+		Control		*P*-value[a]
		n/n	%	*n/n*	%	*n/n*	%	
Meiosis[b]								
GV breakdown	(MI + MII)/(GV + MI + MII)	100/104	96	176/181	97	8963/9130	98	0.163
PB1 extruded	MII/(MI + MII)	84/100	84[e]	161/176	91[e]	8567/8963	96[f]	<0.001
Fertilization[c]								
Fertilized total	≥1PN/inseminated MII[d]	496/592	84	707/826	86	38,785/45,245	86	0.407
2PN formed	2PN/inseminated MII	270/592	46[e]	665/826	81[f]	36,408/45,245	80[f]	<0.0001
Non-2PN formed	Non-2PN/inseminated MII	226/592	38[e]	42/826	5[f]	2,377/45,245	5[f]	<0.0001
Cleavage[c]								
Good-quality blastocyst	Good-quality blastocyst/2PN	19/270	7[e]	124/665	19[f]	7498/36,408	21[f]	<0.0001

[a] Comparison among three groups.

[b] Only oocytes retrieved from follicles measuring 10 mm or wider were included. Degenerated oocytes were excluded.

[c] Oocytes retrieved from both follicles measuring 10 mm or wider and follicles measuring less than 10 mm were included. Degenerated oocytes were not included.

[d] Inseminated metaphase II (MII) includes both genuine MII oocytes and MII oocytes derived from in-vitro maturation.

[e,f] A subgroup whose column ratio is not significantly different at the 0.05 level after adjusting for Fisher's exact probability of multiple comparisons using Hochberg's method.

ACA, anticentromere antibody; ACA+, ACA positive; ANA, antinuclear antibody; ANA+, ANA positive; GV, germinal vesicle; PB1, first polar body; PN, pronuclei.

TABLE 4 DISTRIBUTION OF THE NUMBER OF PRONUCLEUS AND PRONUCLEI

Pronucleus/ pronuclei	ACA+		ANA+		Control		*P*-value[a]
	n	%	*n*	%	*n*	%	
1	8	7[b]	12	3[b,c]	344	2[c]	0.012
2	55	50[b]	349	91[c]	14,055	90[c]	<0.0001
3	14	13[b]	15	4[c]	782	5[c]	0.003
4	13	12[b]	7	2[c]	297	2[c]	<0.0001
5	10	9[b]	1	0[c]	98	1[c]	<0.0001
≥6	10	9[b]	2	1[c]	79	1[c]	<0.0001
Total	110	100	386	100	15,655	100	

[a] Comparison between three groups.

[b,c] A subgroup whose column ratio is not significantly different at the 0.05 level (Fisher's exact probability) after adjusting for multiple comparisons using Hochberg's method. Oocytes achieved from follicles measuring 10 mm or wider were included.

ACA, anticentromere antibody; ANA, antinuclear antibody.

Patient age and anti-Müllerian hormone were also included in the model as known predictors of pregnancy outcome. The analysis revealed that ACA was an independent predictor of live birth, whereas the contribution of other types of ANAs was insignificant (TABLE 5).

Titre-dependent outcomes

Although the number of cases was small, a decreasing trend in the cumulative birth rate was observed for both ACA and ANA titres (Cochran–Armitage trend test: $P < 0.0001$ and $P = 0.0004$, respectively) (**Supplementary Figure 3**).

The effect of a cut-off setting was examined. Lowering the cut-off for antibody positivity to 1:160 did not affect the cumulative live birth rate of ACA+ but increased the cumulative live birth rate of ANA+ to almost the control level (**Supplementary Figure 4**). This is because the subpopulation ($n = 188$) with an ANA titre of 1:160 and a normal live birth rate shifted to the ANA+ group (**Supplementary Figure 2** and **Supplementary Figure 3**). Otherwise, the effects on pregnancy outcome and embryo development were not different from the analysis using the 1:160 cut-off (**Supplementary Table 2**, **Supplementary Table 3** and **Supplementary Table 4**).

DISCUSSION

Summary of findings

The present study showed that ACA is more associated with fewer live births than other ANAs. The decrease seems to be mainly caused by fertilization with abnormal number of pronuclei, failed embryo cleavage and less blastocyst development. Repeated ICSI failure with multiple pronuclei formation and less blastocyst development may be a distinct clinical entity, that should be recognized as specifically attributable to ACA.

Anticentromere antibodies predicts fewer live birth

In the present study, a significant association between ACA and reduced live birth outcome was found. A previous study suggested a reduction in ICSI outcomes achieved using surrogate markers (embryonic and clinical pregnancy rates); however, live birth outcomes, as a true outcome, were never analysed. *Ying et al. (2013b)* reported that the clinical pregnancy rate after ICSI was lower in ACA-positive women (25% [5/20]) than in ANA-negative women (53% [62/116]), although the difference was not statistically significant. The reason that a statistically significant reduction was found may be that our study enrolled approximately five times more participants for screening than that in Ying's study. In addition, the ACA positivity in our study (0.7%) was lower than one-third of that in Ying's study (2.4%), suggesting that we used a higher titer threshold, and thus selected more severely affected patients.

Anticentromere antibodies was the only ANA that predicted live birth outcomes after ICSI. *Fan et al. (2017)* reported that the anti-dsDNA antibody was the ANA with the most significant negative effect on the pregnancy rate of IVF and embryo transfer. In their case-control study, the clinical

TABLE 5 UNIVARIATE AND MULTIVARIATE LOGISTIC REGRESSION ANALYSIS WITH CUMULATIVE LIVE BIRTH AS A DEPENDENT VARIABLE

Clinical measures	Univariate analysis		Multivariate analysis			
			Included variables		Selected variables	
	Odds ratio	95% CI	Odds ratio	95% CI	Odds ratio	95% CI
Age, years	0.84[c]	0.82 to 0.85	0.84[c]	0.83 to 0.86	0.83[c]	0.82 to 0.85
AMH, ng/ml	1.13[c]	1.10 to 1.16	1.03[c]	0.99 to 1.06		
Centromere pattern[a]	0.09[c]	0.02 to 0.39	0.08[c]	0.02 to 0.37	0.08[c]	0.02 to 0.36
Homogenous pattern[a]	0.56[c]	0.31 to 0.99	0.63	0.26 to 1.40		
Speckled pattern[a]	0.58[c]	0.34 to 0.99	0.67	0.32 to 1.40		
Nucleolar pattern[a]	0.92	0.21 to 4.11	1.23	0.27 to 5.70		
Peripheral pattern[b]						

[a] Autoantibody titres were converted to binary values using 1:160 as the cut-off.

[b] No patient showed peripheral pattern.

[c] $P < 0.05$ for univariate or multivariate analysis.

AMH, anti-Müllerian hormone.

pregnancy rate after IVF was the lowest in a patient group with the anti-dsDNA antibody (11.5%) compared with a patient group with any other ANA types (30.2%) and a patient group without any ANA (47.1%). They concluded that the anti-dsDNA antibody is the most significant predictor of conception in infertile women among ANA. Although *Fan et al. (2017)* did not detected ACA specifically, ACA-positive women were likely included in the patient group with other ANA types. If we calculate the clinical pregnancy rate based on their report, the pregnancy rate tends to be lower in women with other ANA (possibly include ACA-positive women) than in women with no ANA (control), although the difference was marginal ($P = 0.0205$, where $P < 0.0167$ is the significant level for multiple comparisons). This suggests that ACA may reduce the pregnancy rate. In this context, applying a multiple logistic regression model to our data, ACA was associated with poor live birth outcomes and that anti-dsDNA antibodies may be confounding; anti-dsDNA antibody, detected as a homogeneous staining pattern, was associated with a live birth outcome in univariate analysis but not multivariate analysis. The presence of the anti-dsDNA antibody, however, was suspected without confirmation in our study, whereas the situation was the opposite in the study by *Fan et al. (2017)*. Further studies are needed that simultaneously identify ACA and anti-dsDNA antibodies.

Multiple pronuclei are pathognomonic for anticentromere antibodies

The present study revealed that one of the most notable and pathognomonic impairments was a change in the number of pronuclei, especially an increase of three or more pronuclei.

In contrast to our observation, researchers who reported the negative effect of ACA found that fertilization with 2PN formation was unaffected; *Ying et al. (2013b)* examined oocytes after ICSI and reported that no ACA or ANA reduced fertilization with 2PN formation. Similarly, *Shirota et al. (2011)* examined oocytes after ICSI and found that fertilization with 2PN formation was the same between patients with ACA and those with other types of ANA, suggesting no ACA-specific effect on fertilization. The investigators (*Hidaka et al., 2015*) subsequently reported that by increasing the number of cases, a reduction in fertilization with 2PN formation was significant for ACA compared with no ACA and other types of ANA. Moreover, they noticed that abnormally fertilized eggs in ACA-positive patients often had multiple pronuclei (≥ 3PN were 35.0%, 0.9%, and 2.4% of the total ICSI oocytes in patients with ACA, other types of ANA, and no ANA, respectively) (*Hidaka et al., 2015*). *Tokoro et al. (2015)* also reported an increase in multiple pronuclei formation, with 51.3% versus 3.3 % of fertilized oocytes in ACA-positive and ACA-negative patients, respectively Consistent with these findings, the present study showed that multiple pronuclei, rarely observed in patients with no ACA, increased up to 50%. Furthermore, relative risk increased markedly with increasing pronuclear number. Therefore, screening for ACA may be recommended when 3PN or more, or more specifically four pronuclei or more, are present (the positive likelihood ratio is calculated to be 5.4 and 10.8 for ACA positivity when 3PN or more and four pronuclei or more were observed, respectively).

Origin of extra pronuclei

Extra pronuclei can be of either female or male origin. Concerning ACA, *Tokoro et al. (2015)* stained pronuclei with an anti-H3K9me2 antibody and found that extra pronuclei were of maternal origin. They also observed that maternal chromosomes were dispersed in the metaphase I oocytes of ACA-positive patients. Therefore, it is possible to suppose that maternal chromosomes are dispersed until pronucleus formation and surrounded by the fusion of nuclear membrane pieces as multiple clusters of chromosomes, which results in fragmented (multiple) pronuclei. It is known that the spindle assembly checkpoint, which detects a spindle-unattached chromosome and blocks the progression to anaphase, is less stringent in the human oocyte, allowing meiotic progression even with erroneous chromosome segregations (*Holubcova et al., 2015*; *Thomas et al., 2021*). This mechanism is similar to that of extra pronuclei of maternal origin in diploid (not triploid) eggs after ICSI (*Grossmann et al., 1997*; *Rosenbusch, 2010*). Our findings of multiple pronucei far beyond 3PN favour the fragmentation origin of multiple pronuclei rather than other mechanisms of 3PN, including the extrusion error of a maternal haploid into the second polar body and non-extrusion of the second polar body, both of which yield a triploid egg (*Van Blerkom et al.; 1984*; *Balakier, 1993*).

Putative action of anticentromere antibodies provided by animal experiments

Animal experiments have shown that ACA impair meiosis of mouse oocytes. Microinjected or medium-derived anticentromere protein A antibody reaches the nucleus, blocks the spindle–kinetochore interaction and interferes with congression, linear arrangement and segregation of chromosomes, leading to segregation error (*Li et al., 2017*; *Liu et al., 2021*). Segregation errors occurs also by microinjection of serum from autoimmune disease patients suspected of having ACAs (*Bernat et al., 1990*; *Simerly et al., 1990*). Moreover, autoantibodies in culture media or in sera can penetrate living cells, including oocytes as is discussed elsewhere (*Zeng et al., 2019*; *Liu et al., 2021*; *Muro et al., 2021*; *Chepy et al., 2022*). Taking these, it is plausible that segregation errors of chromosomes occur in human oocytes retrieved from ACA-positive patients; however, care must be taken when extrapolating from mouse experiments because significant differences exist in the structure and function of the centromere between humans and mice (*Muro et al., 2021*; *Thomas et al., 2021*).

Animal studies suggest that the effect of the antibody is cell cycle dependent. Anticentromere antibodies injected at metaphase or later does not block anaphase progression of the cell cycle, but it does when injected at interphase or earlier (*Bernat et al., 1990*; *Simerly et al., 1990*; *Li et al., 2017*). The 'critical period of the cycle' may explain why meiosis I was less affected, because the critical period (interphase) of meiosis I corresponds to the oogonium period of the foetus, when autoantibodies are not yet produced.

Cleavage failure

The second most notable decline stage observed was embryo cleavage; high-quality blastocyst development from 2PN eggs decreased. The impairment of embryo cleavage seems to be ACA-specific because the impairment was not observed for other types of ANA (TABLE 3).

Li et al. (2017) reported that ACA microinjected into mouse oocytes caused chromosomal dispersion and inhibited homologous chromosome segregation, resulting in aneuploidy. *Holubcova et al. (2015)* examined human oocytes and reported that spontaneously occurring spindle instability causes lagging of

chromosomes and segregation error, resulting in aneuploidy. Aneuploidy may explain both cleavage and implantation failure observed in the present study. Karyotyping of growth-arrested human embryos is needed.

Cumulative outcomes

Per-person outcomes (cumulative outcomes) instead of per-cycle outcomes were analysed. In the present study, a difference was found in background between ACA+ and the other two groups; ACA+ women had more ICSI cycles and stimulation cycles. A woman who finds it difficult to conceive will frequently repeat the ICSI cycles with more stimulation, thereby biasing the per-cycle pregnancy rate and lowering it further. This makes the quantitative comparison more difficult. Therefore, we chose cumulative outcome (per-person outcome) instead of per-cycle outcome so that differences in treatment cycles may not require consideration.

Limitations

This present study has several limitations, including the small number of patients with ACA. Moreover, this study did not consider the coexistence of autoantibodies other than ANA, such as anti-phospholipid antibodies. Therefore, the confounding of other autoantibodies that may coexist with ACA cannot be excluded. Although the comorbidity of collagen diseases was excluded by present symptoms and medical history, patients in the prodromal period of collagen disease could be included. The detection of ACA was based on staining patterns without immunoblotting to specify centromere antigens. Nevertheless, immunostaining diagnosis may be practically reliable for the detection of ACAs because centromere antigens are unique antigens that can be specifically detected by their discrete-speckled pattern, and in turn, the pattern is positive for 93% of progressive systemic sclerosis (*Tan et al., 1980*).

In conclusion, the present study suggests that ACA decreases ICSI outcomes primarily by impairing fertilization and embryonic cleavage. The formation of multiple pronuclei (fragmentation) and recurrent failure to achieve a good-quality blastocyst may be a clue to screening for ANAs. Whether ACA interferes with chromosome segregation in meiosis II by inhibiting centromere function, as the mechanism shown in animal studies, remains to be elucidated.

DATA AVAILABILITY

No data was used for the research described in the article.

ACKNOWLEDGEMENTS

The authors would like to thank Professor Yoshinao Muro, MD, PhD, for his valuable advice on anticentromere antibodies in human sera.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103864.

REFERENCES

Balakier, H., 1993. Tripronuclear human zygotes: the first cell cycle and subsequent development. Hum Reprod 8, 1892–1897.

Bernat, R.L., Borisy, G.G., Rothfield, N.F., Earnshaw, W.C., 1990. Injection of anticentromere antibodies in interphase disrupts events required for chromosome movement at mitosis. J Cell Biol 111, 1519–1533.

Cavalcante, M.B., Cavalcante, C.T.dM.B., Sarno, M., da Silva, A.C.B., Barini, R., 2020. Antinuclear antibodies and recurrent miscarriage: Systematic review and meta-analysis. American Journal of Reproductive Immunology 83, e13215.

Chan, E.K., Damoiseaux, J., Carballo, O.G., Conrad, K., de Melo Cruvinel, W., Francescantonio, P.L., Fritzler, M.J., Garcia-De La Torre, I., Herold, M., Mimori, T., et al., 2015. Report of the first international consensus on standardized nomenclature of antinuclear antibody HEp-2 cell patterns 2014–2015. Front Immunol 6, 412.

Chepy, A., Bourel, L., Koether, V., Launay, D., Dubucquoi, S., Sobanski, V., 2022. Can antinuclear antibodies have a pathogenic role in systemic sclerosis? Front Immunol 13, 930970.

Fan, J., Zhong, Y., Chen, C., 2017. Impacts of anti-dsDNA antibody on in vitro fertilization-embryo transfer and frozen-thawed embryo transfer. J Immunol Res 2017, 8596181.

Grossmann, M., Calafell, J.M., Brandy, N., Vanrell, J.A., Rubio, C., Pellicer, A., Egozcue, J., Vidal, F., Santaló, J., 1997. Origin of tripronucleate zygotes after intracytoplasmic sperm injection. Hum Reprod 12, 2762–2765.

Hidaka, N., Shirota, K., Nagata, Y., Honjou, K., Miyamoto, S., 2015. Effects of anticentromere antiobodies on oocyte maturation and embryo cleavage. Med Bull Fukuoka Univ 42, 81–86.

Holubcova, Z., Blayney, M., Elder, K., Schuh, M., 2015. Human oocytes. Error-prone chromosome-mediated spindle assembly favors chromosome segregation defects in human oocytes. Science 348, 1143–1147.

Kikuchi, K., Shibahara, H., Hirano, Y., Kohno, T., Hirashima, C., Suzuki, T., Takamizawa, S., Suzuki, M., 2003. Antinuclear antibody reduces the pregnancy rate in the first IVF-ET treatment cycle but not the cumulative pregnancy rate without specific medication. Am J Reprod Immunol 50, 363–367.

Li, L., Qi, S.T., Sun, Q.Y., Chen, S.L., 2017. CENP-A regulates chromosome segregation during the first meiosis of mouse oocytes. J Huazhong Univ Sci Technolog Med Sci 37, 313–318.

Li, Y., Wang, Y., Ma, Y., Lan, Y., Jia, C., Liang, Y., Wang, S., 2015. Investigation of the impact of antinuclear antibody on the outcome of in vitro fertilization/intracytoplasmic sperm injection treatment. Taiwan J Obstet Gynecol 54, 742–748.

Liu, H., Zhang, Y., Liu, H., Huang, Q., Ying, Y., 2021. Anticentromere antibody induced by immunization with centromere protein and Freund's complete adjuvant may interfere with mouse early-stage embryo. Reprod Biol Endocrinol 19, 127.

Muro, Y., Yamashita, Y., Koizumi, H., Takeichi, T., Akiyama, M., 2021. Pitfalls in establishing mouse model of female infertility by immunization with human centromere protein. Immunol Lett 239, 20–22.

Rosenbusch, B.E., 2010. A preliminary concept, deduced from cytogenetic analyses, for explaining different types of multipronuclear

oocytes obtained after intracytoplasmic sperm injection. Fertil Steril 94, 2479–2481.
Shirota, K., Nagata, Y., Honjou, K., Tsujioka, H., Yoshizato, T., Miyamoto, S., 2011. Involvement of anticentromere antibody in interference with oocyte meiosis and embryo cleavage. Fertil Steril 95, 2729–2731.
Simerly, C., Balczon, R., Brinkley, B.R., Schatten, G., 1990. Microinjected centromere [corrected] kinetochore antibodies interfere with chromosome movement in meiotic and mitotic mouse oocytes. J Cell Biol 111, 1491–1504.
Simopoulou, M., Sfakianoudis, K., Maziotis, E., Grigoriadis, S., Giannelou, P., Rapani, A., Tsioulou, P., Pantou, A., Kalampokas, T., Vlahos, N., et al., 2019. The impact of autoantibodies on IVF treatment and outcome: a systematic review. Int J Mol Sci 20.
Tan, E.M., Rodnan, G.P., Garcia, I., Moroi, Y., Fritzler, M.J., Peebles, C., 1980. Diversity of antinuclear antibodies in progressive systemic sclerosis. Anti-centromere antibody and its relationship to CREST syndrome. Arthritis Rheum 23, 617–625.
Teramoto, S., Osada, H., Sato, Y., Shozu, M., 2016. Nondominant small follicles are a promising source of mature oocytes in modified natural cycle in vitro fertilization and embryo transfer. Fertil Steril 106, 113–118.
Teramoto, S., Osada, H., Sato, Y., Shozu, M., 2019. Pregnancy and neonatal outcomes of small follicle-derived blastocyst transfer in modified natural cycle in vitro fertilization. Fertil Steril 111, 747–752.
Thomas, C., Cavazza, T., Schuh, M., 2021. Aneuploidy in human eggs: contributions of the meiotic spindle. Biochemical Society Transactions 49, 107–118.
Ticconi, C., Inversetti, A., Logruosso, E., Ghio, M., Casadei, L., Selmi, C., Di Simone, N., 2023. Antinuclear antibodies positivity in women in reproductive age: From infertility to adverse obstetrical outcomes – A meta-analysis. Journal of Reproductive Immunology 155, 103794.
Tokoro, M., Ohno, H., Aoyagi, N., Sonohara, M., Tsuiki, M., Ishihara, K., Funagayama, Y., Kida, Y., Fukunaga, N., Asada, Y., 2015. Multiple pronuclear analysis by immunofluorescence staining in human embryos derived from patients with positive anti-centromere antibodies. Fertii Steril 104, e305.
Van Blerkom, J., Henry, G., Porreco, R., 1984. Preimplantation human embryonic development from polypronuclear eggs after in vitro fertilization. Fertil Steril 41, 686–696.
Ying, Y., Zhong, Y.P., Zhou, C.Q., Xu, Y.W., Miao, B.Y., Wang, Q., Li, J., 2013a. Preliminary investigation of the impact of anticentromere antibody on oocyte maturation and embryo cleavage. Fertil Steril 100, 1585–1589.
Ying, Y., Zhong, Y.P., Zhou, C.Q., Xu, Y.W., Ding, C.H., Wang, Q., Li, J., Shen, X.T., 2013b. A further exploration of the impact of antinuclear antibodies on in vitro fertilization-embryo transfer outcome. Am J Reprod Immunol 70, 221–229.
Ying, Y., Zhong, Y.P., Zhou, C.Q., Xu, Y.W., Wang, Q., Li, J., Shen, X.T., Wu, H.T., 2012. Antinuclear antibodies predicts a poor IVF-ET outcome: impaired egg and embryo development and reduced pregnancy rate. Immunol Invest 41, 458–468.
Zeng, M., Wen, P., Duan, J., 2019. Association of antinuclear antibody with clinical outcome of patients undergoing in vitro fertilization/intracytoplasmic sperm injection treatment: A meta-analysis. Am J Reprod Immunol 82, e13158.
Zhu, Q., Wu, L., Xu, B., Hu, M.H., Tong, X.H., Ji, J.J., Liu, Y.S., 2013. A retrospective study on IVF/ICSI outcome in patients with anti-nuclear antibodies: the effects of prednisone plus low-dose aspirin adjuvant treatment. Reprod Biol Endocrinol 11, 98.

Received 20 September 2023; received in revised form 9 January 2024; accepted 31 January 2024.

RBMO

ELSEVIER

ARTICLE

Total duration of spontaneous blastocyst collapse during the expansion stage is an independent predictor of euploidy and live birth rates

BIOGRAPHY

Jiahong Zhu graduated from Chongqing Medical University, Chongqing (China). She has been engaged in reproductive medicine at Chongqing Health Centre for Women and Children for more than 8 years, mainly responsible for embryo selection and embryo freezing/thawing. Her current research interests include artificial intelligence in reproductive medicine and embryo selection.

Jiahong Zhu[1,2], Jiayi Zou[1,2], Lihong Wu[1,2], Shun Xiong[1,2], Yang Gao[1,2], Junxia Liu[1,2], Guoning Huang[1,2,*], Wei Han[1,2,*]

KEY MESSAGE

The total duration of collapse to re-expansion is a predictor of lower euploidy and live birth rate. Blastocysts with a longer total duration of collapse should be transferred secondary to other blastocysts. When developing algorithms for pregnancy prediction, the duration of spontaneous collapse should be included as a significant variable.

ABSTRACT

Research question: Is the total duration of spontaneous blastocyst collapse to re-expansion before biopsy related to ploidy and live birth rates after single euploid blastocyst transfer?

Design: This was a retrospective cohort study of 600 preimplantation genetic testing cycles for aneuploidy (PGT-A) cycles, involving 2203 biopsied blastocysts, at a large reproductive medicine centre. Features of spontaneous blastocyst collapse from full to expanded stage, before biopsy, were observed using an embryoscope viewer for embryos cultured in a time-lapse incubator. In total, 568 cycles of frozen blastocyst transfers, either single euploid or mosaic, were performed. Correlations between collapse features and PGT-A outcomes were evaluated, as well as live birth rate, following euploid embryo transfer.

Results: Blastocysts with lower morphological quality or delayed development had significantly higher rates of collapse, multiple collapses, and a longer duration of collapse to re-expansion. After controlling for confounders, such as oocyte age, morphological quality of blastocyst, and day of biopsy, multivariate logistic regression revealed that the total duration of collapse to re-expansion was an independent predictor of lower euploidy rate; the multivariate OR was 0.85 (95% CI 0.77–0.95; $P = 0.00$). Furthermore, even with euploid embryo transfer, the probability of a live birth decreased as the total duration of collapse to re-expansion increased; the multivariate OR was 0.79 (95% CI 0.64–0.98; $P = 0.033$).

Conclusion: The total duration of blastocyst collapse to re-expansion could be used as a predictor of lower euploidy and live birth rate. When developing blastocyst algorithms for pregnancy prediction, the duration of spontaneous blastocyst collapse should be included as a significant variable.

[1] Chongqing Clinical Research Centre for Reproductive Medicine, Chongqing Health Centre for Women and Children, Chongqing, China
[2] Chongqing Key Laboratory of Human Embryo Engineering, Centre for Reproductive Medicine, Women and Children's Hospital of Chongqing Medical University, Chongqing, China

*Corresponding authors. E-mail addresses: gnhuang217@sina.com (G. Huang), hw2012ivf@sina.com (W. Han). https://doi.org/10.1016/j.rbmo.2024.103863 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS

Number of collapses
Total duration of collapse to re-expansion
Euploidy
Embryo quality

INTRODUCTION

Decades after the development of assisted reproductive technology (ART), the live birth rate has improved considerably. Currently, a main goal of ART clinics is to reduce the risk of multiple pregnancies through single embryo transfer. Optimizing ART to select embryos with the greatest implantation potential and maintaining an adequate success rate remains a challenge.

Embryo quality, including chromosomal status and viability, has always been considered an important predictor of successful implantation in ART. Aneuploidy, present in approximately 50% of embryos throughout preimplantation development, contributes significantly to failed implantation (*Rabinowitz et al., 2012*). Frozen euploid blastocyst transfer after preimplantation genetic testing for aneuploidy (PGT-A) has shown improved outcomes, especially for women aged >37 years (*Kang et al., 2016*). However, *Fragouli et al. (2011)* found that nearly one-third of high-quality euploid blastocysts failed to implant, suggesting the need for additional markers to predict embryonic competence and reduce the time for a successful pregnancy when multiple euploid blastocysts are available after PGT-A.

Compared with conventional embryo morphological assessment, the time lapse technique maintains stable culture conditions, and provides continual recording of morphological and morphokinetic parameters that are not included in the Gardner and Schoolcraft scoring system (*Basile et al., 2014*, *2015*; *Liu et al., 2020*). Previous studies have suggested that spontaneous collapse of the blastocyst is a normal physiological phenomenon during the expansion stage, as the blastocyst breaks through the zona pellucida to achieve hatching (*Gonzales et al., 1996*). However, this view has been overturned since the introduction of the time lapse technique. Studies have shown that spontaneous collapse is not necessary for hatching, and can be a pathological phenomenon. Collapsed blastocysts are less likely to be euploid (*Cimadomo et al., 2022*; *Gazzo et al., 2020*) and have lower implantation potential (*Marcos et al., 2015*; *Sciorio et al., 2020*). However, the results are still controversial. *Bodri et al. (2016)* reported that blastocyst collapse was not an independent predictor of reduced live birth rate after adjusting for confounders.

A comprehensive review and meta-analysis by *Bickendorf et al. (2023)* found that spontaneous collapse has negative implications for pregnancy, but the subgroup analyses based on various definitions of spontaneous collapse and the number of collapses included in the analysis yielded inconclusive results. In addition, morphological quality and day of biopsy of the blastocyst have been shown to be associated with the presence of spontaneous collapse and the number of collapses (*Cimadomo et al., 2022*). Therefore, it is possible that the morphological quality of blastocysts and the timeframe of development have significant confounding impacts on the association between spontaneous characteristics, ploidy and pregnancy outcomes. The characteristics of blastocyst collapse are complex, involving a different number of spontaneous collapses and different durations of collapse to re-expansion. Therefore, neither the occurrence of collapse nor its frequency can accurately describe the properties of a spontaneously collapsed blastocyst. Spontaneous collapse of a blastocyst is associated with excessive energy consumption, and has an adverse effect on subsequent development (*Biggers et al., 1998*; *Marcos et al., 2015*; *Sciorio et al., 2021*). On the basis of the above findings, it was postulated that the longer the total duration of collapse to re-expansion in the process of blastocyst expansion, the more energy will be consumed, and the duration of collapse may provide a new indicator to predict blastocyst aneuploidy and pregnancy outcomes in PGT-A cycles.

The aim of this study was to add knowledge to the existing literature by determining whether the number of collapses and total duration of blastocyst collapse to re-expansion can be used to predict euploidy and live birth.

METHODS

Study population

This retrospective study was conducted at the Centre for Reproductive Medicine of the Women and Children's Hospital of Chongqing Medical University from June 2019 to December 2021. In total, 600 PGT-A cycles with 2203 blastocysts were included in this study, with 38 patients having two PGT-A cycles. All embryos were cultured in time-lapse incubators (EmbryoScope+; Vitrolife, Sweden). The number of collapses and total duration of collapse during blastocyst expansion were observed using an embryoscope viewer. In total, 568 frozen blastocyst transfer cycles, either single euploid or mosaic, were performed.

All patients signed consent forms for all the procedures during PGT-A treatment. This study was approved by the Institutional Review Board of Chongqing Health Centre for Women and Children (Reference No. 2020-RGI-05, approval date 22 June 2020).

The primary outcomes were the associations between collapse features and euploidy rates. The secondary outcomes were the associations between collapse features and clinical outcomes.

Ovarian stimulation, insemination and blastocyst culture

The patients underwent long gonadotrophin-releasing hormone (GnRH) agonist or GnRH antagonist protocols for controlled ovarian stimulation, depending on their ovarian response and medical history of ART treatment (*Ding et al., 2021*). Human chorionic gonadotrophin (HCG) (Ovidrel; Merck Serono, Italy) was administered when at least three dominant follicles with an average diameter of >18 mm were obtained. Oocyte retrieval was performed vaginally under ultrasound monitoring 36 h after HCG administration.

Intracytoplasmic sperm injection (ICSI) was performed for all PGT-A cycles. Oocyte denudation was performed 38 h after HCG administration, followed by ICSI 2 h later. After microinjection, all oocytes were cultured individually in an EmbryoScope+ incubator under 6% CO_2, 5% O_2 and 37°C. Embryos were cultured using pre-equilibrated G-1 culture medium (Vitrolife) until 66–68 h after fertilization; then the medium was replaced with pre-equilibrated G-2 medium (Vitrolife), and embryos were cultured until the day of trophectoderm biopsy.

Blastocyst morphological evaluation, biopsy and collapse observations

Before biopsy, blastocyst morphology was evaluated using the Gardner and Schoolcraft scoring system (*Schoolcraft et al., 1999*). The trophectoderm biopsy procedure was performed on the day of full blastocyst expansion, and the inner cell mass and trophectoderm were not both graded as C. A laser (Hamilton Torne Inc., USA) was used to open a 15–20-μm hole in the zona pellucida. Next, five to 10

trophectoderm cells were aspirated gently for next-generation sequencing analysis, as recommended by ESHRE PGT Consortium and SIG-Embryology Biopsy Working Group (*ESHRE PGT Consortium and SIG-Embryology Biopsy Working Group et al., 2020*). After biopsy, blastocysts were vitrified using a Cryotop and Kitazato Vitrification Kit (Kitazato BioPharma Co., Japan) in accordance with the manufacturer's instructions. In brief, vitrification was performed at room temperature (23–25 ℃), with embryos initially suspended in an equilibration solution for 9~15 min, followed by exposure to the vitrification solution for 45–60 s. Finally, embryos were loaded on to a Cryotop tip, and plunged into liquid nitrogen immediately.

The PGT-A outcomes were categorized as euploid, low-level mosaic (<50% aneuploid cells), high-level mosaic (≥50% aneuploid cells) and aneuploid. According to the morphological grading, the biopsied blastocysts were classified into four groups: top quality (AA), good quality (AB and BA), fair quality (BB) and poor quality (AC, CA, BC and CB).

In the current study, fresh blastocysts were cultured in an EmbryoScope+ incubator and observed using Embryoviewer. Spontaneous collapse was defined as the phenomenon in which the trophectoderm surface separates from the zona pellucida, with a separation percentage >50%. The separation was measured using the drawing tool in Embryoviewer. All spontaneous collapses for the blastocyst from full to expanded stage, before trophectoderm biopsy, were included. The duration of spontaneous collapse was defined as the period from the beginning of the collapse until all trophectoderm cells were back in contact with the zona pellucida of the blastocyst. The total duration of collapse for a blastocyst was estimated by summing the duration of each collapse (FIGURE 1).

Frozen embryo transfer and pregnancy diagnosis

Euploid embryos were transferred individually in a hormonal replacement treatment cycle as described previously (*Ye et al., 2019*). Frozen embryo transfer (FET) can be performed when the endometrial thickness reaches ≥7 mm. The blastocysts were warmed in accordance with the manufacturer's guidelines using the Kitazato Vitrification Kit (Kitazato BioPharma Co.). Following warming, the blastocysts were transferred to G2 medium (G2; VitroLife Sweden AB) and cultured at 37°C and 6% CO_2 in an air incubator. FET was performed 1–2 h after the euploid blastocyst was warmed. The luteal phase was supported with a combination of oestrogen (Progynova; Bayer, Germany) and progesterone (Baiyunshan, China) until 12 weeks of gestation. Serum β-HCG levels were measured 14 days after FET to determine pregnancy. Twenty-eight days after FET, patients who tested positive for HCG (>5 mIU/ml) underwent a vaginal ultrasound examination. Implantation was defined as the presence of an intrauterine gestational sac with a visible fetal heartbeat detected during transvaginal ultrasound. Miscarriage was defined as an intrauterine pregnancy loss before 28 weeks of gestation. Live birth was defined as the delivery of one or more living infants at a gestation of >28 weeks.

Statistical analysis

Continuous quantitative variables are presented as mean ± SD or median [interquartile range (IQR)], and categorical data are expressed as a percentage. Comparisons between groups were conducted using Student's *t*-test (for normally distributed variables), or Kruskal–Wallis rank sum test and Wilcoxon–Mann–Whitney test (for non-normally distributed variables) for continuous variables, and chi-squared test for categorical data. Generalized estimating equation multivariate logistic regression was performed to confirm significant associations, adjusting for oocyte age, morphological quality of blastocyst, and day of biopsy. A significance level of $P < 0.05$ was used for all statistical tests. All analyses were performed using R (http://www.r-project.org) and EmpowerStats software (www.

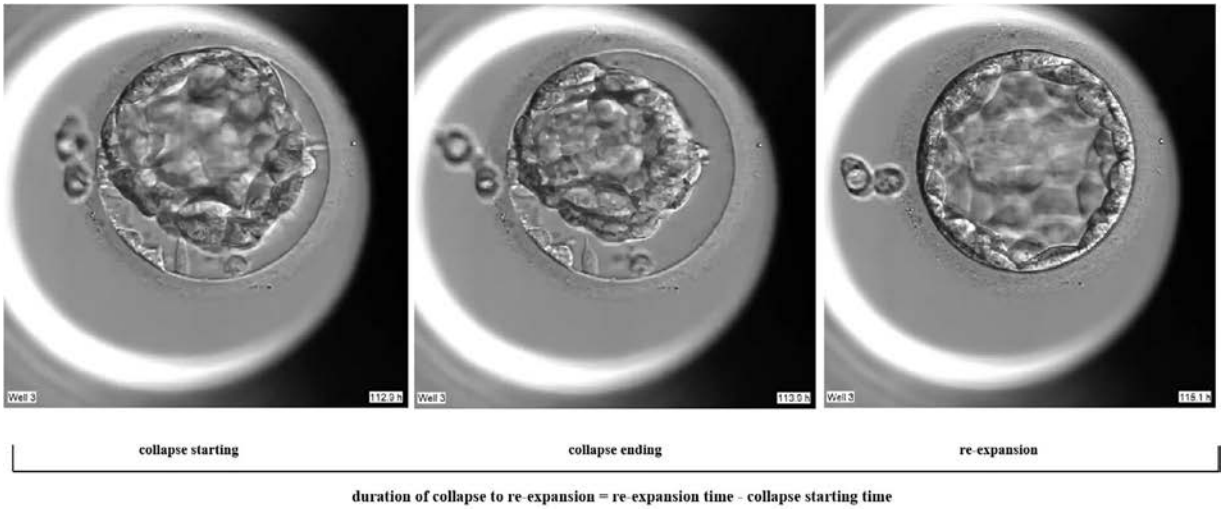


FIGURE 1 Definition of the duration of collapse to re-expansion. Spontaneous collapse was defined as the phenomenon in which the trophectoderm surface separates from the zona pellucida, with a separation percentage >50%. The duration of spontaneous collapse was defined as the period from the beginning of the collapse until all trophectoderm cells were back in contact with the zona pellucida of the blastocyst. The formula for the duration of collapse to re-expansion is the re-expansion time minus the collapse start time. Pictured example shows duration of collapse to re-expansion: 2.2 h.

empowerstats.com; X&Y Solutions, Inc., USA).

RESULTS

In the current study, 2203 blastocysts from 600 PGT-A cycles were biopsied and analysed. Of the 600 cycles, 130 (21.67%) had no collapsing blastocysts, and 470 (78.33%) had at least one collapsing blastocyst. Based on the presence or absence of spontaneous blastocyst collapse, the cycles were divided into two groups: cycles with no collapsing blastocysts, and cycles with at least one collapsing blastocyst. The baseline patient characteristics of the two groups are listed in TABLE 1. Significant differences in previous IVF attempts (1.99 ± 1.98 versus 1.62 ± 1.34; $P = 0.013$), anti-Müllerian hormone (2.90 ± 2.12 versus 3.61 ± 2.55; $P=0.005$), days of gonadotrophin administration (8.30 ± 2.19 versus 8.70 ± 1.77; $P = 0.031$), and number of oocytes retrieved (9.98 ± 6.09 versus 13.40 ± 7.00; $P < 0.001$) were found between the two groups. Multivariate logistic regression was used to determine whether the patient characteristics listed in TABLE 1 were associated with the likelihood of cycles which involved at least one collapsing blastocyst. The findings showed no significant associations (Supplementary Table 1).

Features of spontaneous blastocyst collapse and PGT-A outcomes

All of the 2203 biopsied blastocysts were processed for genetic testing. FIGURE 2 depicts the association between the PGT-A results and characteristics of spontaneous blastocyst collapse [presence or absence of collapse(s), and number of collapses]. Overall, euploidy rates were significantly lower in spontaneously collapsing blastocysts compared with those in non-collapsing blastocysts (39.52% versus 46.42%, respectively; $P = 0.002$; FIGURE 2A), and aneuploidy rates were significantly higher in spontaneously collapsing blastocysts compared with non-collapsing blastocysts (48.87% versus 43.40%, respectively; $P = 0.015$; FIGURE 2A). However, there was no discernible difference in the mosaicism rate between the two groups. Furthermore, the number of spontaneous collapses was negatively correlated with euploidy rate. Euploidy rates were 42.94%, 35.43% and 26.92% for blastocysts with one, two and three or more collapses, respectively ($P = 0.001$) (FIGURE 2B). For the collapsed embryos, the median total durations of collapse to re-expansion were 1.60 (IQR 1.10–2.60), 2.20 (IQR 1.20–3.05), 1.80 (IQR 1.18–2.65) and 1.95 (IQR 1.3–3.2) h for euploid, low-level mosaic, high-level mosaic and aneuploid blastocysts, respectively (FIGURE 3). There were significant differences in the total duration of collapse to re-expansion between euploid and both aneuploid and low-level mosaic embryos ($P < 0.001$ for euploid versus aneuploid; $P = 0.002$ for euploid versus low-level mosaic) (FIGURE 3).

TABLE 1 BASELINE CHARACTERISTICS OF CYCLES WITH OR WITHOUT SPONTANEOUS BLASTOCYST COLLAPSE

Variable	Cycles without collapsing blastocysts (*n* = 130)	Cycles with at least one collapsing blastocyst (n = 470)	*P*-value
Female age (years)	33.08 ± 4.89	32.31 ± 4.75	0.102
Male age (years)	35.15 ± 6.94	34.14 ± 5.63	0.087
Previous IVF attempts	1.99 ± 1.98	1.62 ± 1.34	0.013
Duration of infertility (years)	3.40 ± 3.79	2.86 ± 3.47	0.126
Number of previous miscarriages	0.96 ± 1.35	1.03 ± 1.37	0.637
Type of infertility			0.917
Primary infertility	36 (27.69%)	128 (27.23%)	
Secondary infertility	94 (72.31%)	342 (72.77%)	
Aetiology of infertility			0.820
Tubal factor	49 (37.69%)	193 (41.06%)	
Male factor	23 (17.69%)	67 (14.26%)	
Endometriosis	2 (1.54%)	10 (2.13%)	
Ovulation disorders	0 (0.00%)	4 (0.85%)	
Combination	8 (6.15%)	30 (6.38%)	
Unexplained infertility	48 (36.92%)	166 (35.32%)	
Basal AMH (ng/ml)	2.90 ± 2.12	3.61 ± 2.55	0.005
Basal FSH (IU/l)	6.08 ± 2.23	5.89 ± 3.08	0.559
Basal LH (IU/l)	3.25 ± 1.34	3.62 ± 2.44	0.132
Female BMI (kg/m^2)	21.76 ± 2.64	22.20 ± 2.88	0.121
Gn (days)	8.30 ± 2.19	8.70 ± 1.77	0.031
Total dose of Gn used (IU)	1918.95 ± 762.14	1892.97 ± 639.77	0.697
Number of oocytes retrieved	9.98 ± 6.09	13.40 ± 7.00	<0.001

Data shown are mean ± SD or *n* (%).

Independent-sample Student's *t*-tests and chi-squared test were used for comparisons.

Repeat patients had measures recorded for each cycle.

AMH, anti-Müllerian hormone; BMI, body mass index; Gn, gonadotrophin.

Features of spontaneous blastocyst collapse and morphological quality

The embryo collapse rate was negatively related to morphological grade, with rates of 22.94%, 24.17%, 35.17% and 55.95% for top, good, fair and poor quality, respectively ($P < 0.001$) (TABLE 2). Additionally, the rate of multiple collapses increased with decreasing blastocyst quality (all $P = 0.000$) (TABLE 2). The total durations of collapse to re-expansion were 2.04 ± 1.35, 1.91 ± 1.38, 2.24 ± 1.56 and 2.95 ± 1.95 h for top-, good-, fair- and poor-quality blastocysts, respectively (TABLE 2). Top-quality and good-quality blastocysts had a similar total duration of collapse to re-expansion, while for the other groups, the total duration of collapse to re-expansion showed an increasing trend as the morphological grade decreased. For blastocysts with poor morphological quality, the total duration of collapse to re-expansion was significantly longer than that of the other morphological quality groups (compared with top quality: $P = 0.004$; compared with good quality: $P < 0.001$; compared with fair quality: $P < 0.001$).

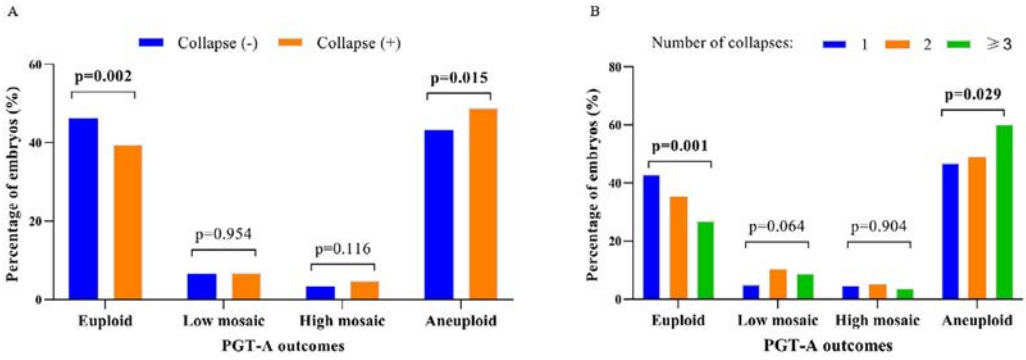


FIGURE 2 Relationship between the preimplantation genetic testing cycles for aneuploidy (PGT-A) outcomes of biopsied blastocysts and features of spontaneous collapse. (A) PGT-A outcomes according to the presence or absence of collapse(s). (B) PGT-A outcomes according to the number of collapses. Statistical significance was assessed through chi-squared test. All significant *P*-values are presented in bold.

Features of spontaneous blastocyst collapse and day of biopsy

Blastocysts that were biopsied on day 5 displayed significantly lower rates of collapse (24.28% versus 50.43%, respectively; $P < 0.001$; TABLE 3), were less likely to have multiple collapses (TABLE 3), and had a shorter total duration of collapse to re-expansion (1.98 ± 1.32 versus 2.61 ± 1.84 h, respectively; $P < 0.001$; TABLE 3) compared with embryos biopsied on day 6.

Multivariate logistic regression analysis for the association between blastocyst euploidy and collapse features

In this study, 971 blastocysts were euploid, 997 were aneuploid, 149 were low-level mosaic, and 86 were high-level mosaic. Multivariate logistic regression was used to analyse the relationship between blastocyst euploidy and collapse features (TABLE 4). After controlling for oocyte age, morphological quality of blastocyst, and day of biopsy, the results showed that the number of collapses had no significant impact on euploidy rate (each $P > 0.05$), whereas the total duration of collapse to re-expansion was significantly related to blastocyst euploidy rate; the multivariate OR was 0.85 (95% CI 0.77–0.95; $P = 0.000$).

Features of spontaneous blastocyst collapse and pregnancy outcomes

In total, 568 single frozen blastocyst transfer (sFET) cycles were performed, of which 530 were euploid and 38 were mosaic. For the 38 sFET cycles using mosaic blastocysts, 15 did not result in pregnancy, four resulted in miscarriage, and 19 resulted in live birth. The associations between the collapse features and pregnancy outcomes of mosaic blastocysts were not analysed statistically because of the limited number. For the 530 sFET cycles using euploid blastocysts, 173 did not result in pregnancy, 47 resulted in miscarriage, and 310 resulted in live birth. The characteristics of pregnancy outcomes of transferred euploid blastocysts are summarized in TABLE 5. The euploid sFET cycles had a significantly higher ratio of good-quality blastocysts, a lower ratio of poor-quality blastocysts, and a shorter total duration of collapse to re-expansion in blastocysts that resulted in a live birth compared with those that did not ($P = 0.007$, $P=0.027$ and $P =0.015$, respectively). There were no discernible differences between day of biopsy and rate of collapse.

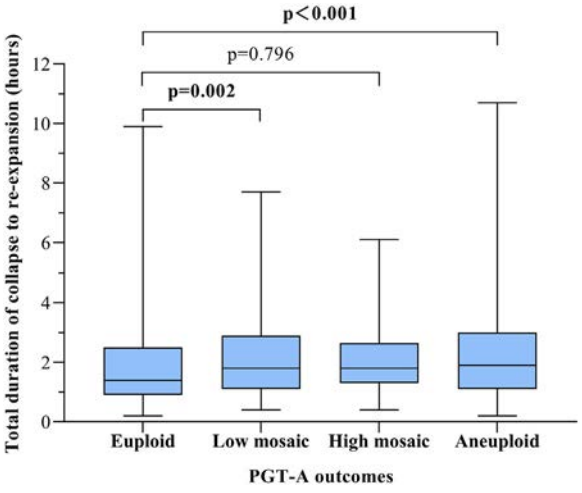


FIGURE 3 Total duration of blastocyst collapse to re-expansion by different preimplantation genetic testing cycles for aneuploidy (PGT-A) outcomes. The total duration of collapse by PGT-A outcomes was displayed using a box plot. The median and interquartile range (IQR) are represented by box and whiskers. Kruskal–Wallis tests revealed a significant overall *P*-value of 0.001 among the four groups. Wilcoxon–Mann–Whitney tests were conducted to evaluate the differences between euploid embryos and those classified as aneuploid, low-level mosaic and high-level mosaic. All significant *P*-values are presented in bold.

Multivariate logistic regression was used to analyse the associations between the features of euploid blastocysts before transfer and the likelihood of a live birth (TABLE 6). The results showed that the total duration of collapse to re-expansion was inversely correlated with the live birth rate after adjusting for variables (oocyte age, female BMI, day of biopsy, and morphological quality of blastocyst). The multivariate OR was 0.79 (95% CI 0.64–0.98; $P = 0.033$). However, the number of collapses had no significant effect on the live birth rate.

DISCUSSION

The purpose of this study was to evaluate the relationships between characteristics

TABLE 2 DESCRIPTION OF COLLAPSE FEATURES ACCORDING TO THE MORPHOLOGICAL QUALITY OF BLASTOCYSTS

Collapse features	Top quality (*n* = 170)	Good quality (*n* = 633)	Fair quality (*n* = 1089)	Poor quality (*n* = 311)	Top quality versus poor quality *P*-value	Good quality versus poor quality *P*-value	Fair quality versus poor quality *P*-value
Total duration of collapse to re-expansion (h)	2.04 ± 1.35	1.91 ± 1.38	2.24 ± 1.56	2.95 ± 1.95	0.004	<0.001	<0.001
Collapsed embryo rate (%)	39 (22.94%)	153 (24.17%)	383 (35.17%)	174 (55.95%)	<0.001	<0.001	<0.001
Number of collapses							
0	131 (77.06%)	480 (75.83%)	706 (64.83%)	137 (44.05%)	<0.001	<0.001	<0.001
1	29 (17.06%)	127 (20.06%)	261 (23.97%)	79 (25.40%)	0.036	0.035	0.326
2	8 (4.71%)	17 (2.69%)	92 (8.45%)	58 (18.65%)	0.000	0.000	0.000
≥3	2 (1.18%)	9 (1.42%)	30 (2.75%)	37 (11.90%)	0.000	0.000	0.000

Data are shown as *n* (%) and mean ± SD.

Fisher's exact, Kruskal–Wallis and chi-squared tests were used to determine statistical significance.

Morphological quality of blastocysts was graded according to the Gardner and Schoolcraft scoring system. Blastocysts with grades of AA were assigned to the top-quality group; BA and AB to the good-quality group; BB to the fair-quality group; and AC, CA, BC and CB to the poor-quality group.

of biopsied blastocyst collapse, blastocyst ploidy and live birth after single euploid blastocyst transfer. After controlling for confounders (oocyte age, female BMI, morphological quality of blastocyst, and day of biopsy), multivariate logistic regression revealed that the total duration of collapse to re-expansion was significantly associated with the euploidy rate, and negatively correlated with the live birth rate after single euploid blastocyst transfer.

Embryonic development capacity is influenced significantly by patient physiology. However, the findings of this study indicated that there were no correlations between patient or cycle characteristics and the likelihood that at least one embryo within a cohort would collapse. Recent studies of the effects of collapsed blastocysts on pregnancy outcomes had controversial results. According to *Marcos et al. (2015)*, spontaneous collapse reduced the implantation potential of blastocysts, and transferring collapsed blastocysts when there are other embryos available was not advised. Similar findings were reported by *Sciorio et al. (2020)*, who showed that spontaneously collapsing blastocysts had lower pregnancy rates than non-collapsing embryos. In contrast, a study by *Bodri et al. (2016)* showed that collapse was not an independent predictor of live birth after adjusting for confounders, such as female age and morphokinetic variables.

In terms of the association between collapse features and blastocyst euploidy, *Gazzo et al. (2020)* reported that the presence of blastocyst collapse was associated with ploidy status, and collapsed embryos were more likely to be chromosomally abnormal, regardless of female age. This result was supported by a recent study by *Cimadomo et al. (2022)*, who defined spontaneous collapse as an event involving an uninterrupted reduction in the area of the zona pellucida that occurs after the initiation of blastulation. This process lasts for <10 h, and the ratio of the collapsed embryo area to the zona pellucida area at the end of the process was ≤90%. Their study carefully documented the parameters of spontaneous blastocyst collapse, such as the area of the blastocyst at the time of biopsy, number of collapses, timing of biopsy, and duration of each collapse (time between the start and end times of collapse). *Cimadomo et al. (2022)* revealed that only the occurrence of spontaneous blastocyst collapse was an effective indicator of aneuploidy. Euploidy was not substantially correlated with the frequency of subsequent collapses, length of the longest collapse, size of the biggest shrinkage, or ratio of the collapsed embryo area to the zona pellucida area at the end of collapse.

TABLE 3 DESCRIPTION OF COLLAPSE FEATURES ACCORDING TO DAY OF BIOPSY

Collapse features	Day 5 (*n* = 1384)	Day 6 (*n* = 819)	*P*-value
Total duration of collapse to re-expansion (h)	1.98 ± 1.32	2.61 ± 1.84	<0.001
Collapsed embryo rate	336 (24.28%)	413 (50.43%)	<0.001
Number of collapses			
0	1048 (75.72%)	406 (49.57%)	<0.001
1	260 (18.79%)	236 (28.82%)	0.000
2	61 (4.41%)	114 (13.92%)	0.000
≥3	15 (1.08%)	63 (7.69%)	0.000

Data are presented as mean ± SD and *n* (%).

Statistical significance was assessed using chi-squared and Kruskal–Wallis tests.

The current findings demonstrated a negative correlation between the presence of spontaneous collapse and the euploidy rate. The blastocysts that had poorer morphological quality displayed significantly greater rates of collapse, multiple collapses, and a longer duration of collapse to re-expansion. Additionally, a previous study demonstrated that low-quality embryos have a lower rate of euploidy (*Li et al., 2022*). The present results showed that blastocysts biopsied on day 6 displayed significantly greater rates of collapse, multiple collapses, and a longer duration of collapse to re-expansion. Previous research also showed that a

TABLE 4 MULTIVARIATE LOGISTIC REGRESSION FOR THE ASSOCIATION BETWEEN BLASTOCYST EUPLOIDY AND BLASTOCYST CHARACTERISTICS

Model effect	Multivariate OR (95% CI)	Adjusted *P*-value
Blastocyst quality		
Top	Control	
Good	0.69 (0.50–0.95)	0.02
Fair	0.45 (0.33–0.62)	0.00
Poor	0.37 (0.25–0.55)	0.00
Day of biopsy		
5	Control	
6	0.96 (0.79–1.18)	0.72
Number of collapses		
0	Control	
1	1.21 (0.93–1.58)	0.16
2	1.21 (0.76–1.92)	0.42
≥3	1.17 (0.53–2.6)	0.70
Total duration of collapse to re-expansion (h)	0.85 (0.77–0.95)	0.00

Multivariate logistic regression was used to estimate the associations between euploidy and blastocyst features, controlling for oocyte age, morphological quality of blastocyst, and day of biopsy.

Morphological quality of blastocysts was graded according to the Gardner and Schoolcraft scoring system. Blastocysts with grades of AA were assigned to the top-quality group; BA and AB to the good-quality group; BB to the fair-quality group; and AC, CA, BC and CB to the poor-quality group.

higher number of spontaneous collapses was associated with delayed blastocyst expansion (*Cimadomo et al., 2022*).

However, it is noteworthy that in the current study, blastocysts cultured to days 5 and 6 had different durations of monitoring, and thus the higher number of collapses in day 6 blastocysts could be the result of the longer monitoring period. In addition, embryos with delayed development are more likely to have chromosomal abnormalities (*Bamford et al., 2022*). Therefore, the present study used multivariate logistic regression to account for confounding variables, which included patient age, morphological quality of blastocyst, and day of biopsy. The results showed no significant correlation between collapse frequency and euploidy; however, the total duration of collapse to re-expansion was highly correlated with euploidy.

A recent study on the association between implantation potential and collapse characteristics of euploid blastocysts revealed a decreasing trend in the live birth rate across blastocysts grouped based on an increasing number of collapses (*Cimadomo et al., 2022*). These results are consistent with a previous study by *Viñals Gonzalez et al. (2018)*, which showed that blastocyst contraction behaviour compromised reproductive competence in high-quality euploid embryos. Interestingly, in the current study, multivariate logistic regression showed that the live birth rate decreased when one euploid blastocyst with a longer total

TABLE 5 TRANSFERRED EUPLOID BLASTOCYST CHARACTERISTICS BY PREGNANCY OUTCOME

Variable	Live birth (+) (*n* = 310)	Live birth (−) (*n* = 220)			Live birth (+)versus live birth (−) *P*-value
		Did not implant (*n* = 173)	Miscarriage (*n* = 47)	Total (*n* = 220)	
Blastocyst quality					
Top	30 (9.68%)	20 (11.56%)	5 (10.64%)	25 (11.36%)	0.531
Good	110 (35.48%)	46 (26.59%)	8 (17.02%)	54 (24.55%)	0.007
Fair	146 (47.10%)	85 (49.13%)	26 (55.32%)	111 (50.45%)	0.446
Poor	24 (7.74%)	22 (12.72%)	8 (17.02%)	30 (13.64%)	0.027
Collapse rate (%)	93 (30.00%)	57 (32.95%)	16 (34.04%)	73 (33.18%)	0.436
Day of biopsy					
5	238 (76.77%)	118 (68.21%)	39 (82.98%)	157 (71.36%)	0.159
6	72 (23.23%)	55 (31.79%)	8 (17.02%)	63 (28.64%)	0.159
Number of collapses					
0	217 (70.00%)	116 (67.05%)	31 (65.96%)	147 (66.82%)	0.436
1	74 (23.87%)	40 (23.12%)	11 (23.40%)	51 (23.18%)	0.854
2	14 (4.52%)	12 (6.94%)	5 (10.64%)	17 (7.73%)	0.121
≥3	5 (1.61%)	5 (2.89%)	0 (0.00%)	5 (2.27%)	0.748
Total duration of collapse to re-expansion (h)	0.78 ± 1.21	1.09±1.48	0.93±1.23	1.06 ± 1.43	0.015

Data are shown as *n* (%) and mean ± SD.

Fisher's exact, Kruskal–Wallis and chi-squared tests were used to determine statistical significance.

Morphological quality of blastocysts was graded according to the Gardner and Schoolcraft scoring system. Blastocysts with grades of AA were assigned to the top-quality group; BA and AB to the good-quality group; BB to the fair-quality group; and AC, CA, BC and CB to the poor-quality group.

TABLE 6 MULTIVARIATE LOGISTIC REGRESSION FOR THE ASSOCIATION BETWEEN LIVE BIRTH AND EUPLOID BLASTOCYST CHARACTERISTICS

Model effect	Live birth versus non-live birth	
	Adjusted OR (95% CI)	*P*-value
Blastocyst quality		
Top	Control	–
Good	1.66 (0.88–3.13)	0.118
Fair	1.17 (0.64–2.15)	0.608
Poor	0.75 (0.33–1.68)	0.484
Number of collapses		
0	Control	–
≥1	1.06 (0.53–2.10)	0.868
Day of biopsy	1.54 (0.86–2.76)	0.145
Day 5	Control	–
Day 6	1.00 (0.64–1.54)	0.985
Total duration of collapse to re-expansion (h)	0.79 (0.64–0.98)	0.033

Multivariate logistic regression was used to estimate the associations between live birth and blastocyst features, controlling for oocyte age, female BMI, morphological quality of blastocyst, and day of biopsy.

Morphological quality of blastocysts was graded according to the Gardner and Schoolcraft scoring system. Blastocysts with grades of AA were assigned to the top-quality group; BA and AB to the good-quality group; BB to the fair-quality group; and AC, CA, BC and CB to the poor-quality group.

duration of collapse to re-expansion was transferred, after controlling for potential confounders including female age, number of collapses, morphological quality of blastocyst, and day of biopsy.

It is known that blastocyst expansion during embryonic development depends on the Na+/K+-ATPase pump, which increases the osmotic pressure to draw water into the embryo. This increase in fluid causes the blastocoel to expand, the zona pellucida to thin, and the blastocyst to hatch (*Baltz et al., 1997*; *Le Verge-Serandour et al., 2021*). Currently, there are no robust explanations for blastocyst collapse and energy consumption. *Viñals Gonzalez et al. (2018)* showed that chromosome content plays a crucial role in early embryogenesis, and it seems to be closely related to some features, such as embryo collapse. Previous studies proposed that the blastocysts showing spontaneous collapse and re-expansion were attempting to rescue a state of relatively disorganized growth of the trophectoderm or, alternatively, were in the process of extruding damaged (possible aneuploid) cells (*Coticchio et al., 2021*). In addition, *Kobayashi et al. (2020)* showed that the abundance of cell-free mitochondrial DNA in spent medium was associated with blastocyst collapse, and they speculated that cell-free mitochondrial DNA may be produced by dead or fragmented blastomeres of blastocysts, arise from specific events in the blastocyst, or are actively secreted by blastocysts. The findings of the current study indicate that there is strong correlation between the duration of collapse to re-expansion and the likelihood of achieving a live birth with euploid embryo transfer. It was observed that the chance of obtaining euploid embryos was significantly decreased as the total duration of collapse to re-expansion increased. Furthermore, even with euploid embryo transfer, the probability of achieving a live birth decreased with an increasing duration of collapse to re-expansion.

The reduced live birth rate with an increased duration of collapse to re-expansion could be explained by the fact that extruding damaged cells and subsequently re-expanding the blastocyst requires a lot of energy. As the duration of collapse to re-expansion is prolonged, this energy consumption becomes more pronounced, which could have a negative impact on the viability of the embryo. As a result, the probability of a live birth decreases as the duration of collapse to re-expansion increases. These findings highlight the importance of considering the duration of collapse to re-expansion during the blastocyst expansion stage.

A strength of the current study is the consideration of potential confounding factors that could affect the results, such as morphological quality of blastocyst and day of biopsy. Additionally, this study provided detailed information on features of spontaneous collapse, not merely stating the presence of collapse. Furthermore, all embryos were biopsied at the fully expanded stage, ensuring consistency in the analysis and evaluation of collapse.

A limitation of this study is its retrospective nature. The rate of spontaneous collapse, rate of multiple collapses, and total duration of collapse to re-expansion may also vary with different laboratory environments, culture systems and patient characteristics. Previous studies showed that the number of spontaneous collapses was associated with morphological quality of blastocyst; the worse the blastocyst quality, the longer the duration of collapse (*Cimadomo et al., 2022*). However, only biopsied embryos were included, so the ploidy of embryos that did not meet biopsy criteria was unknown. Further well-designed studies are needed to confirm the current results, and elucidate the role of collapse events during blastocyst expansion in predicting pregnancy outcomes.

CONCLUSIONS

This study showed that the total duration of spontaneous blastocyst collapse to re-expansion during the expansion stage could be used as a predictor of aneuploidy and live birth rate in PGT-A cycles. Blastocysts with a longer total duration of collapse should be used as a secondary option for transfer. When developing blastocyst algorithms for pregnancy prediction, the duration of spontaneous blastocyst collapse should be included as a significant variable.

DATA AVAILABILITY

Data will be made available on request.

ACKNOWLEDGEMENTS

The authors would like to thank the colleagues who helped with the study and the data collection. In addition, the authors wish to thank Liwen Bianji (Edanz) (www.liwenbianji.cn/) for editing the English text of a draft of this manuscript.

FUNDING

This work was supported by the Natural Science Foundation Project of Chongqing, Chongqing Science and Technology Commission (Grant Nos. CSTB2023NSCQ-MSX0258; and CSTB2022NSCQ-MSX0253).

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103863.

REFERENCES

Baltz, J.M., Smith, S.S., Biggers, J.D., Lechene, C., 1997. Intracellular ion concentrations and their maintenance by Na+/K(+)-ATPase in preimplantation mouse embryos. Zygote 5, 1–9 PMID:9223240.

Bamford, T., Barrie, A., Montgomery, S., Dhillon-Smith, R., Campbell, A., Easter, C., Coomarasamy, A., 2022. Morphological and morphokinetic associations with aneuploidy: a systematic review and meta-analysis. Hum Reprod Update 28, 656–686. https://doi.org/10.1093/humupd/dmac022.

Basile, N., Nogales, M.delC., Bronet, F., Florensa, M., Riqueiros, M., Rodrigo, L., García-Velasco, J., Meseguer, M., 2014. Increasing the probability of selecting chromosomally normal embryos by time-lapse morphokinetics analysis. Fertil Steril 101, 699–704. https://doi.org/10.1016/j.fertnstert.2013.12.005.

Basile, N., Vime, P., Florensa, M., Aparicio Ruiz, B., García Velasco, J. A., Remohí, J., Meseguer, M, 2015. The use of morphokinetics as a predictor of implantation: a multicentric study to define and validate an algorithm for embryo selection. Hum Reprod 30, 276–283. https://doi.org/10.1093/humrep/deu331.

Bickendorf, K, Qi, F, Peirce, K, Natalwala, J, Chapple, V, Liu, Y., 2023. Spontaneous collapse as a prognostic marker for human blastocysts: a systematic review and meta-analysis. Hum Reprod 38, 1891–1900. https://doi.org/10.1093/humrep/dead166.

Biggers, J.D., 1998. Reflections on the culture of the preimplantation embryo. Int J Dev Biol 42, 879–884 PMID:9853817.

Bodri, D., Sugimoto, T., Yao Serna, J., Kawachiya, S., Kato, R., Matsumoto, T., 2016. Blastocyst collapse is not an independent predictor of reduced live birth: a time-lapse study. Fertil Steril 105, 1476–1483.e3. https://doi.org/10.1016/j.fertnstert.2016.02.014.

Cimadomo, D., Marconetto, A., Trio, S., Chiappetta, V., Innocenti, F., Albricci, L., Erlich, I., Ben-Meir, A., Har-Vardi, I., Kantor, B., Sakov, A., Coticchio, G., Borini, A., Ubaldi, F.M., Rienzi, L., 2022. Human blastocyst spontaneous collapse is associated with worse morphological quality and higher degeneration and aneuploidy rates: a comprehensive analysis standardized through artificial intelligence. Hum Reprod 37, 2291–2306. https://doi.org/10.1093/humrep/deac175.

Coticchio, G., Barrie, A., Lagalla, C., Borini, A., Fishel, S., Griffin, D., Campbell, A., 2021. Plasticity of the human preimplantation embryo: developmental dogmas, variations on themes and self-correction. Hum Reprod Update 27, 848–865. https://doi.org/10.1093/humupd/dmab016.

Ding, X., Yang, J., Li, L., Yang, N., Lan, L., Huang, G., Ye, H., 2021. Fertility outcomes in women after controlled ovarian stimulation with gonadotropin releasing hormone agonist long protocol: fresh versus frozen embryo transfer. BMC Pregnancy Childbirth 21, 207. https://doi.org/10.1186/s12884-021-03698-5.

Consortium, ESHRE PGT, Group, SIG-Embryology Biopsy Working, Kokkali, G., Coticchio, G., Bronet, F., Celebi, C., Cimadomo, D., Goossens, V., Liss, J., Nunes, S., Sfontouris, I., Vermeulen, N., Zakharova, E., De Rycke, M., 2020. ESHRE PGT Consortium and SIG Embryology good practice recommendations for polar body and embryo biopsy for PGT. Hum Reprod Open 2020 (3), hoaa020. https://doi.org/10.1093/hropen/hoaa020.

Fragouli, E., Alfarawati, S., Daphnis, D.D., Goodall, N.N., Mania, A., Griffiths, T., Gordon, A., Wells, D., 2011. Cytogenetic analysis of human blastocysts with the use of FISH, CGH and aCGH: scientific data and technical evaluation. Hum Reprod 26, 480–490. https://doi.org/10.1093/humrep/deq344.

Gazzo, E., Peña, F., Valdéz, F., Chung, A., Velit, M., Ascenzo, M., Escudero, E., 2020. Blastocyst contractions are strongly related with aneuploidy, lower implantation rates, and slow-cleaving embryos: a time lapse study. JBRA Assist Reprod 24, 77–81. https://doi.org/10.5935/1518-0557.20190053.

Gonzales, D.S., Jones, J.M., Pinyopummintr, T., Carnevale, E.M., Ginther, O.J., Shapiro, S.S., Bavister, B.D., 1996. Trophectoderm projections: a potential means for locomotion, attachment and implantation of bovine, equine and human blastocysts. Hum Reprod 11, 2739–2745. https://doi.org/10.1093/oxfordjournals.humrep.a019201.

Kang, H.J., Melnick, A.P., Stewart, J.D., Xu, K., Rosenwaks, Z., 2016. Preimplantation genetic screening: who benefits? Fertil Steril 106, 597–602. https://doi.org/10.1016/j.fertnstert.2016.04.027.

Kobayashi, M., Kobayashi, J., Shirasuna, K., Iwata, H., 2020. Abundance of cell-free mitochondrial DNA in spent culture medium associated with morphokinetics and blastocyst collapse of expanded blastocysts. Reprod Med Biol 19, 404–414. https://doi.org/10.1002/rmb2.12344.

Le Verge-Serandour, M., Turlier, H, 2021. A hydro-osmotic coarsening theory of biological cavity formation. PLoS Comput Biol 17, e1009333. https://doi.org/10.1371/journal.pcbi.1009333.

Li, N., Guan, Y., Ren, B., Zhang, Y., Du, Y., Kong, H., Zhang, Y., Lou, H., 2022. Effect of Blastocyst Morphology and Developmental Rate on Euploidy and Live Birth Rates in Preimplantation Genetic Testing for Aneuploidy Cycles With Single-Embryo Transfer. Front Endocrinol (Lausanne) 13, 858042. https://doi.org/10.3389/fendo.2022.858042.

Liu, Y., Qi, F., Matson, P., Morbeck, D.E., Mol, B.W., Zhao, S., Afnan, M., 2020. Between-laboratory reproducibility of time-lapse embryo selection using qualitative and quantitative parameters: a systematic review and meta-analysis. J Assist Reprod Genet 37, 1295–1302. https://doi.org/10.1007/s10815-020-01789-4.

Marcos, J., Pérez-Albalá, S., Mifsud, A., Molla, M., Landeras, J., Meseguer, M., 2015. Collapse of blastocysts is strongly related to lower implantation success: a time-lapse study. Hum Reprod 30, 2501–2508. https://doi.org/10.1093/humrep/dev216.

Rabinowitz, M., Ryan, A., Gemelos, G., Hill, M., Baner, J., Cinnioglu, C., Banjevic, M., Potter, D., Petrov, D.A., Demko, Z., 2012. Origins and rates of aneuploidy in human blastomeres. Fertil Steril 97, 395–401. https://doi.org/10.1016/j.fertnstert.2011.11.034.

Schoolcraft, W.B., Gardner, D.K., Lane, M., Schlenker, T., Hamilton, F., Meldrum, D.R., 1999. Blastocyst culture and transfer: analysis of results and parameters affecting outcome in two in vitro fertilization programs. Fertil Steril 72, 604–609. https://doi.org/10.1016/s0015-0282(99)00311-8.

Sciorio, R., Herrer Saura, R., Thong, K.J., Esbert Algam, M., Pickering, S.J., Meseguer, M.,

2020. Blastocyst collapse as an embryo marker of low implantation potential: a time-lapse multicentre study. Zygote 1–9. https://doi.org/10.1017/S0967199419000819.

Sciorio, R., Meseguer, M., 2021. Focus on time-lapse analysis: blastocyst collapse and morphometric assessment as new features of embryo viability. Reprod Biomed Online 43, 821–832. https://doi.org/10.1016/j.rbmo.2021.08.008.

Viñals Gonzalez, X., Odia, R., Cawood, S., Gaunt, M., Saab, W., Seshadri, S., Serhal, P., 2018. Contraction behaviour reduces embryo competence in high-quality euploid blastocysts. J Assist Reprod Genet 35, 1509–1517. https://doi.org/10.1007/s10815-018-1246-x.

Ye, H., Luo, X., Pei, L., Li, F., Li, C., Chen, Y., Zhang, X., Huang, G., 2019. The addition of single dose GnRH agonist to luteal phase support in artificial cycle frozen embryo transfer: a randomized clinical trial. Gynecol Endocrinol 35, 618–622. https://doi.org/10.1080/09513590.2018.1563888.

Received 19 October 2023; received in revised form 26 January 2024; accepted 30 January 2024.

RBMO

ELSEVIER

ARTICLE

Artificial intelligence-powered assisted ranking of sibling embryos to increase first cycle pregnancy rate

BIOGRAPHY

Dr Boussommier-Calleja is a biomedical engineer trained in EPFL (Switzerland), Imperial College London (UK) and MIT (USA) and has co-authored more than 10 peer-reviewed papers. She has received funding from Marie Curie Individual Fellowship, the Irvington Fellowship and from the Cancer Research Institute, among others.

Nina Dissler[a], Daniela Nogueira[b,c], Bertrand Keppi[d], Pierre Sanguinet[e], Christophe Ozanon[f], Cendrine Geoffroy-Siraudin[g], Xavier Pollet-Villard[h], Alexandra Boussommier-Calleja[a,*]

KEY MESSAGE

EMBRYOLY artificial intelligence score correlates with clinical pregnancies and live births. Use of its ranking alongside the expertise of embryologists within cohorts of sibling embryos could have increased first cycle pregnancy rate and reduced time to pregnancy.

ABSTRACT

Research question: Could EMBRYOLY, an artificial intelligence embryo evaluation tool, assist embryologists to increase first cycle pregnancy rate and reduce cycles to pregnancy for patients?

Design: Data from 11,988 embryos were collected via EMBRYOLY from 2666 egg retrievals (2019–2022) across 11 centres in France, Spain and Morocco using three time-lapse systems (TLS). Data from two independent clinics were also examined. EMBRYOLY's transformer-based model was applied to transferred embryos to evaluate ranking performances against pregnancy and birth outcomes. It was applied to cohorts to rank sibling embryos (including non-transferred) according to their likelihood of clinical pregnancy and to compute the agreement with the embryologist's highest ranked embryo. Its effect on time to pregnancy and first cycle pregnancy rate was evaluated on cohorts with multiple single blastocyst transfers, assuming the embryologist would have considered EMBRYOLY's ranking on the embryos favoured for transfer.

Results: EMBRYOLY's score correlated significantly with clinical pregnancies and live births for cleavage and blastocyst transfers. This held true for clinical pregnancies from blastocyst transfers in two independent clinics. In cases of multiple single embryo transfers, embryologists achieved a 19.8% first cycle pregnancy rate, which could have been improved to 44.1% with the adjunctive use of EMBRYOLY (McNemar's test: $P < 0.001$). This could have reduced cycles to clinical pregnancy from 2.01 to 1.66 (Wilcoxon test: $P < 0.001$).

[a] ImVitro, Paris, France, 130 Rue de Lourmel, 75015 Paris.
[b] INOVIE Fertilité, Institut de Fertilité La Croix Du Sud, Clinique la Croix du Sud, Toulouse, France.
[c] Art Fertility Clinics, IVF laboratory, Abu Dhabi, United Arab Emirates.
[d] INOVIE Group, INOVIE Fertilié, Gen-Bio, 63100 Clermont-Ferrand, France.
[e] INOVIE Group, INOVIE Fertilié, LaboSud, 34000 Montpellier, France.
[f] Cerballiance, 69005, Lyon, France.
[g] Hopital Saint Joseph, Service Médecine et biologie de la reproduction, 13285, Marseille, France.
[h] MLAB Groupe, Centre d'Assistance Médicale à la Procréation Nataliance, Pôle Santé Oréliance, Saran, France.

*Corresponding author. E-mail address: alexandra@im-vitro.com (A. Boussommier-Calleja). https://doi.org/10.1016/j.rbmo.2024.103887 1472-6483/
Declaration: AB-C is a co-owner of, and holds stocks in, ImVitro SAS. AB-C holds a patent for 'Devices and processes for machine learning prediction of in vitro fertilization' (EP20305914.2). ND is an employee of ImVitro and has been granted stock options. XP-V. has been granted stock options of ImVitro. The remaining authors have no conflicts to declare.

KEYWORDS

artificial intelligence
embryo evaluation
time to pregnancy
pregnancy
birth
time-lapse incubator systems

Conclusions: EMBRYOLY's potential to enhance first cycle pregnancy rates when combined with embryologists' expertise is highlighted. It reduces the number of unsuccessful cycles for patients across TLS and IVF centres.

INTRODUCTION

Embryo evaluation remains a core task for embryologists and consists predominantly of ranking embryos according to their morphology and morphokinetics (*Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011*) to transfer embryos with the highest chance of pregnancy first, hoping to minimize the number of failed transfers and reduce time to pregnancy for patients. The first report characterizing a deep learning algorithm aiming at evaluating embryos was published by *Tran et al. (2019)*. This algorithm was, however, predominantly trained on discarded embryos (*Kan-Tor et al., 2020*; *Tran et al., 2020*), and only on recordings derived from one type of time-lapse incubator. Since then, more algorithms have been characterized to assist embryologists in reducing their time spent evaluating embryos and to minimize intra- and inter-operator variability, in the hopes that improved and standardized embryo evaluation may increase the likelihood of selecting high-quality embryos, potentially shortening time to pregnancy and increasing first cycle pregnancy rates. To date, however, the only study that has investigated the effect of artificial intelligence models on time to pregnancy in real cohorts of sibling embryos accomplished it by binary comparison (increase versus decrease), without quantifying the difference in number of cycles (*Cimadomo et al., 2023*).

Such algorithms have been trained using a combination of any of the following options: on single images of blastocysts (*VerMilyea et al., 2020*) or on series of images recorded by time-lapse systems (TLS) to exploit the rich information from embryo kinetics; to predict an outcome, e.g. ploidy (*Chavez-Badiola et al., 2020*) or blastulation (*Petersen et al., 2016*; *Liao et al., 2021*), or to recognize a feature already visible on the images, e.g. number of pronuclei (*Fukunaga et al., 2020*) or cleavage state (*Zabari et al., 2023*); on an objective end point, e.g. pregnancy or live birth, or on a subjective criteria of success, e.g. embryo quality (*Chen et al., 2019*; *Khosravi et al., 2019*; *Kragh et al., 2019*); on visual information alone (*Theilgaard Lassen et al., 2023*) or on hybrid datasets that also combined categorical clinical features for more accurate predictions (*Miyagi et al., 2020*; *Barnes et al., 2023*; *Duval et al., 2023*).

Evaluating and comparing such artificial intelligence algorithms is complex (*Kragh and Karstoft, 2021*), and is often boiled down to reporting their area under the curve computed across a large validation set of embryos. This metric is equivalent to the probability of two randomly selected embryos being correctly ranked, but it falls short of illustrating how the model would have affected the embryologist's decision in the real-life condition of a cohort of sibling embryos. Such algorithms have been trained to recognize a binary outcome, e.g. pregnancy or no pregnancy, which does not guarantee that they can relatively rank sibling embryos. To the best of our knowledge, only two recent studies have evaluated deep learning models in the context of cohorts of embryos. For example, *Diakiw et al. (2022)* evaluated time to pregnancy of a computer vision tool trained on static images of blastocysts, and *Cimadomo et al. (2023)* evaluated iDAScore v1.0, an algorithm trained to predict clinical pregnancy based on time-lapse recording from the Embryoscope® or Embryoscope+® (Vitrolife, Västra Frölunda, Sweden). The former, however, evaluated their algorithm on fictive cohorts of embryos without being able to compare it to the ranking of the embryologists. The latter evaluated the model using cohorts of embryos from a single centre that had systematically been pre-screened for preimplantation genetic testing for aneuploidy (PGT-A); however, this fails to represent the workflow of many IVF centres that do not carry out PGT-A. Their study, however, went beyond pregnancy and attempted to measure the effect on time to live birth. Importantly, both studies assessed an artificial intelligence model used as standalone, which would not consider any input from the embryologist. While this is an interesting approach, it evaluates an artificial intelligence tool that replaces the embryologist, which is an unlikely use of any artificial intelligence tool today. More realistically, an embryologist would want to participate in the ranking of embryos to some extent, and not blindly follow an artificial intelligence tool; the synergy between an artificial intelligence tool and the ranking of the embryologist has, to the best of our knowledge, never been evaluated. Evaluating an artificial intelligence tool as standalone, however, is not just problematic because it is unrealistic; it also supposes that the artificial intelligence tool needs to consider all embryos from a cohort, many of which have not been transferred. This is why *Cimadomo et al. (2023)* could not quantify the time to live birth of iDAScore v1.0 as standalone, as it had often ranked an embryo that was never transferred as the highest. Although artificial intelligence models for embryo evaluation will undoubtedly assist with IVF automation and potentially be used as a standalone in the future, a complete transition to solely artificial intelligence-assisted embryo evaluation remains questionable, as the clinical experience of the embryologist has significant value (*Fitz et al., 2021*). The use of artificial intelligence embryo evaluation tools as they become routine for embryologists is more likely going to be complementary to embryologists, which has not yet been evaluated.

The aim of the present study was to evaluate whether EMBRYOLY (*Duval et al., 2023*), together with the initial ranking of the embryologist, could have improved outcomes in cohorts of embryos. The described findings of this multi-centric and multi-TLS retrospective study suggest that using EMBRYOLY as an additional datapoint could have improved the first cycle clinical pregnancy rate and reduced the number of cycles to clinical pregnancy, consolidating its potential added value in clinically relevant laboratory settings.

MATERIALS AND METHODS

Data description

The present study used retrospective data uploaded through EMBRYOLY from 11 clinics in France, Spain and Morocco, corresponding to fresh and frozen IVF cycles completed between 2019 and 2022. Data from two of the 11 clinics had never been used for training the algorithm. This study used retrospective and de-identified data uploaded through EMBRYOLY by IVF centres. Collection of retrospective data for this study was exempted from ethical review and approval, and from the requirement for informed consent,

because of the retrospective nature of the analyses and de-identification of data.

The data consisted of videos of embryos recorded using one of the following time-lapse systems: Embryoscope® or Embryoscope+® (Vitrolife, Västra Frölunda, Sweden), GERI® (Genea BiomedX, Rowville, Australia) or MIRI® (ESCO Medical, Egå, Denmark). Data from a total of 11988 embryos were collected from 2666 oocyte retrievals. Some patients may have had more than one oocyte retrieval; all oocyte retrievals are considered separately. Details are presented in Supplementary Table 1. Embryologists could add directly on EMBRYOLY the following information per embryo: whether it was discarded, frozen, transferred fresh or transferred frozen; the date of transfer if transferred; its transfer outcome if known at the time of data entry; and information about the patients and their treatment (or donor when relevant). Clinical variables (Supplementary Table 1) were reported at the patient level at the time of egg retrieval; they did not vary from one embryo to the other for a given cohort. The transfer outcomes were categorized as follows: clinical pregnancy if a fetal heartbeat via ultrasound was detected between 6 and 9 weeks of amenorrhoea (FH+); and if no heartbeat was detected between 6 and 9 weeks of amenorrhoea (FH−); live birth, if, after 22 completed weeks of gestational age, breathing, or any other evidence of life, was present and birth weight was at least 500 g. Transfer dates were also added by embryologists, which helped confirm which embryos were transferred fresh and which ones were transferred frozen subsequently. Only one centre (corresponding to 33% of the data) carried out oocyte warming regularly. In such cases, our ranking was evaluated in the context of the separate batches of thawed oocytes.

Algorithm description

The algorithm evaluated in this study is part of the EMBRYOLY device. A modified version of the algorithm described in *Duval et al. (2023)* was used in this study with comparable performances. In *Duval et al. (2023)*, a 3D ResNet was used; here a UniFormer was used (*Li et al., 2022*), which is a transformer-based architecture specially designed for video processing (*Vaswani et al., 2017*). This new model integrates three-dimensional convolution in the shallow layers and self-attention in the deep layers in order to tackle spatio−temporal dependencies with a suitable balance between computation and accuracy. The model is made of 21 million parameters and was pretrained on an open-source dataset named Kinetics. The latter is an open-source video dataset provided by DeepMind that covers 400 human action classes. This type of model has the advantage of being particularly powerful to classify videos in an efficient way. The model takes as its input a series of images from the central focal plane of the TLS or, from the plane considered to have the highest contrast by the TLS when available at the time of data collection, i.e. GERI. The target of this network is to predict the fetal heartbeat likelihood from its input video (FH+/FH−) by outputting a score between 0 (lower likelihood of clinical pregnancy) and 1 (higher likelihood of clinical pregnancy), a score that will be evaluated using the binary cross entropy. Note that only the computer vision layer of the algorithm described in *Duval et al. (2023)* was used, without taking into account any feature describing the patients and their treatment. This is because the present study aimed to evaluate the ranking of embryos, which would not have been affectedby clinical characteristics that are common to all embryos within the same cohort.

Experimental design

Three experiments were carried out on three different subsets of data. First, to test whether EMBRYOLY's score can rank embryos as a function of their clinical pregnancy or live birth potential as a support tool for embryologists, all embryos with known transfer outcome were evaluated together (n = 2657 considering clinical pregnancy and 2121 considering live birth) with a logistic regression, including transfers at the cleavage stage (n = 579 considering clinical pregnancy and 475 considering live birth) and blastocyst stage (n = 2078 considering clinical pregnancy and 1646 considering live birth).

Two clinic hold-out analyses were conducted by testing the logistic regression solely on embryos from two IVF centres in which data had never been used to train the algorithm. The first held-out centre is a French clinic that carried out 210 oocyte retrievals, including 256 embryos with known clinical pregnancy outcome. The second is a Moroccan clinic that carried out 63 oocyte retrievals, including 90 embryos with known clinical pregnancy outcome.

Second, to test how often the embryo ranked highest by EMBRYOLY as a standalone was also the one ranked highest by the embryologist, only cohorts that had several embryos available for transfer of which only one was transferred on days 5−7 was considered (n = 926 oocyte retrievals). All embryos of the cohorts were evaluated (n = 4974 embryos), including those that were not transferred. Embryologists assessed embryos using a mix of morphological and kinetic criteria, which could differ between clinics and practitioners. The agreement consisted of computing the number of cohorts where the embryo transferred was the same as the one ranked highest by EMBRYOLY. If two embryos from a cohort had the exact same EMBRYOLY score, they were both considered as the highest ranked; this only happened in 4% of the 926 egg retrievals considered.

Third, to test whether EMBRYOLY's ranking combined with the expertise of embryologists would have minimized the number of cycles required to reach a clinical pregnancy (CTP), or increased the first cycle clinical pregnancy rate (FCP), only cohorts of embryos with multiple single embryo transfers at the blastocyst stage, including one transfer that led to a clinical pregnancy and one that did not, were analysed (n = 111 oocyte retrievals, n = 260 embryos). To compute the CTP and FCP, only embryos transferred were considered to have their transfer outcome (FIGURE 1). This analysis rested on two main hypotheses: the embryo transferred first by the embryologist was the one they ranked highest; the transfer outcome would have been the same had the embryos been transferred fresh or frozen and in a different order for a given patient. The embryologist CTP was computed for each egg retrieval as the number of transfer cycles that had been needed to reach a clinical pregnancy based on their own decision (following their own standard assessment criteria), including fresh and frozen transfers.

To illustrate how embryologists would have been supported by EMBRYOLY, its ranking was only considered on embryos that had been transferred by embryologists and thus deemed most promising in their initial evaluation; the embryo transferred first by the embryologist was prioritized in case it obtained the same EMBRYOLY score as another embryo from the cohort (which only happened in 2% of the 111 egg retrievals). The CTP of embryologists

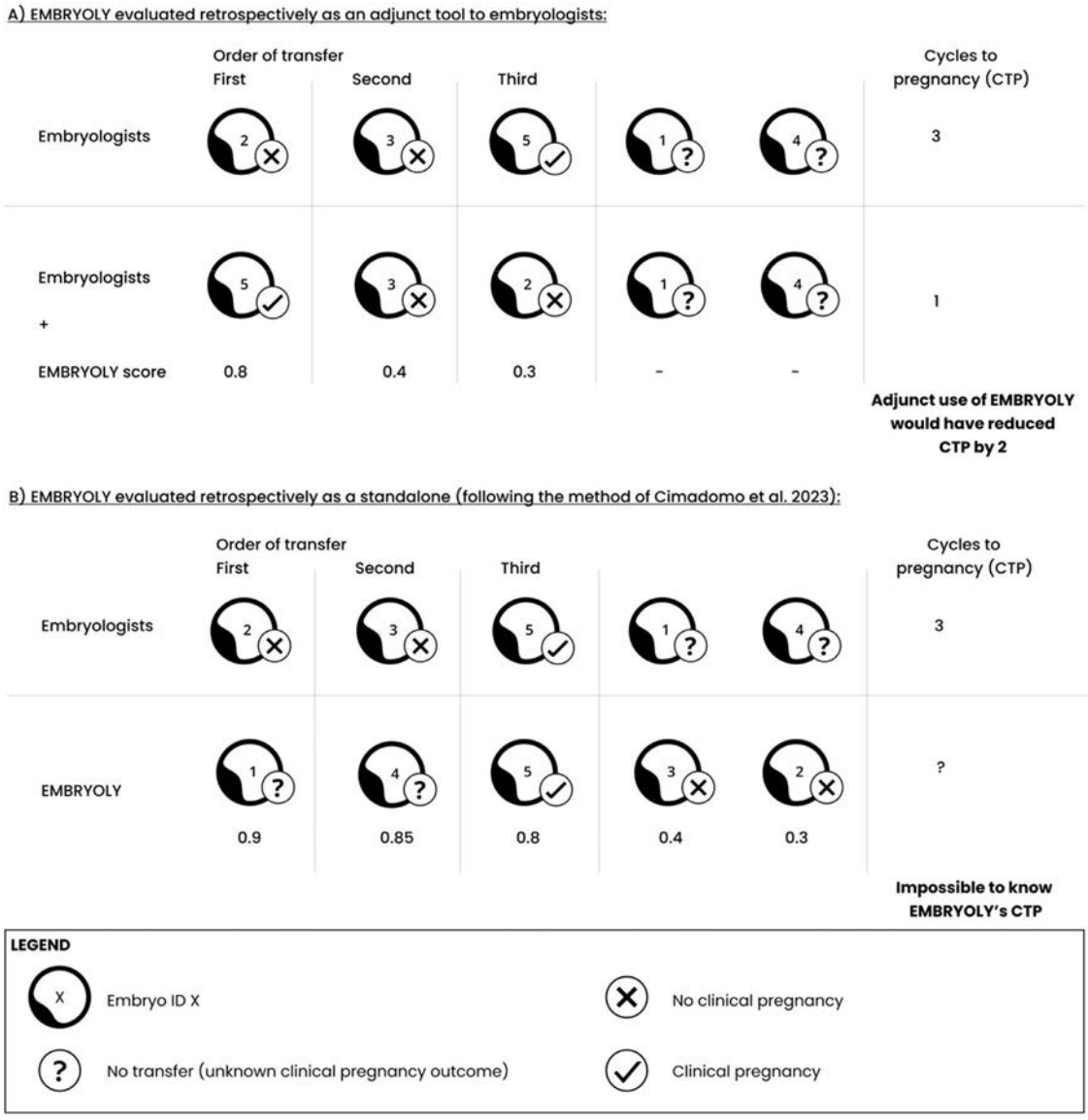


FIGURE 1 Experimental retrospective design that (A) evaluates EMBRYOLY as a software adjunct to embryologists, compared with (B) the retrospective design of the evaluation of EMBRYOLY as a standalone, i.e. evaluating the model applied directly to cohorts without the need for preselection or interpretation by embryologists, suggested by *Cimadomo* et al. *(2023)*. In this example, a cohort that contains five embryos, including three single embryo transfers, is considered. The embryologist had transferred embryo 2 first, which did not lead to a clinical pregnancy, embryo 3 as a second transfer, which did not lead to a clinical pregnancy either and embryo 5 as a third transfer, which did lead to a clinical pregnancy.
(A) The embryos that were not transferred are not considered because it is presumed that EMBRYOLY would be used on the initial pre-selection of the embryologist. If the embryologist had seen EMBRYOLY's scores for these three embryos, we see in this example that embryo 5 would have been transferred first because it was scored highest by EMBRYOLY and, therefore, only one cycle would have been needed to reach a clinical pregnancy, compared with three for the embryologist alone; (B) If EMBRYOLY is assumed to automate the decision making as a standalone, we consider scores for all embryos including those that were not transferred and thus favoured by embryologists. Here, embryos 1 and 4 would have been ranked highest; however, as they were not transferred, it is impossible to quantify the effect of EMBRYOLY on the number of cycles to clinical pregnancy. CTP, cycles to pregnancy; ID,identifier.

supported by EMBRYOLY, therefore, consisted of counting the number of cycles required to reach a clinical pregnancy when considering EMBRYOLY's ranking on the embryologist's initial ranking (FIGURE 1). The FCP of embryologists supported by EMBRYOLY was computed as the percentage of egg retrievals in which the transferred embryo ranked highest by EMBRYOLY would have led to a clinical pregnancy and compared with the actual FCP resulting from the embryos the embryologist decided to transfer first.

Statistics

Logistic regressions were used to assess statistical associations between EMBRYOLY's score and transfer outcome on day-3 and day-5–7 embryos: odds ratios with 95% confidence intervals and *P*-values are reported. The agreement between embryologists and EMBRYOLY for embryo selection is reported as a percentage of agreement with a standard error. The simulation of the effect of EMBRYOLY on FCP was supported by a McNemar's test to compare FCP of embryologists alone and of embryologists with the support of EMBRYOLY. The

simulation of the effect of the adjunct use of EMBRYOLY on CTP was supported by a Wilcoxon test to compare CTP of embryologists alone and of embryologists with the support of EMBRYOLY. Across all statistical tests, a significance threshold of 5% was used.

RESULTS

Patients included in this study were on average 34.1 ± 4.6 years old (ranging from 19–45 years), with an average body mass index of 24.1 ± 4.5 (Supplementary Table 1).

EMBRYOLY score used adjunctively to embryologists can statistically rank embryos transferred at the cleavage or blastocyst stage as a function of their likelihood of clinical pregnancy and live birth outcomes

A logistic regression analysis conducted on embryos transferred at the blastocyst stage showed that an increase of +0.01 in EMBRYOLY's score was statistically associated with +2.8% relative increase in likelihood of clinical pregnancy (OR 1.028, 95% CI 1.023 to 1.033, $P < 0.001$; $n = 2078$ transferred blastocysts), demonstrating that EMBRYOLY can rank blastocysts according to their clinical pregnancy potential (FIGURE 2). Of the nine clinics that presented a sufficient sample size ($n > 35$ blastocysts), the logistic regression with clinical pregnancy remained statistically significant for seven clinics (Supplementary Table 2), including the two held-out clinics whose data have never been used in the training of the algorithm (French clinic: OR 1.022, 95% CI 1.003 to 1.042; $P = 0.02$; $n = 220$ transferred blastocysts; Moroccan clinic: OR 1.026, 95% CI 1.003 to 1.051, $P = 0.03$; $n = 79$ transferred blastocysts).

The logistic regression on embryos transferred at the blastocyst stage was also statistically significant when looking at live birth (OR 1.024, 95% CI 1.016 to 1.032, $P < 0.001$; $n = 1646$ transferred blastocysts). Even when only data from individual clinics was considered, this conclusion held true in five out of eight clinics eligible for analysis (Supplementary Table 2), including the Moroccan hold-out clinic (OR 1.032, 95% CI 1.002 to 1.063; $P = 0.04$; $n = 56$ transferred blastocysts).

Finally, when embryos transferred on day 3 were evaluated, the logistic regression was statistically significant on clinical pregnancy (OR 1.026, 95% CI 1.017 to 1.036, $P < 0.001$; $n = 579$ transferred cleavage embryos) as well as on live birth (OR = 1.019, 95% CI 1.004 to 1.035, $P = 0.01$; $n = 475$ transferred cleavage embryos) (FIGURE 2). Neither analyses (clinical pregnancy and live birth after cleavage embryo transfer) could be applied on the held-out clinics because the sample size was too small ($n = 34$ and 11 transferred cleavage embryos with known birth outcome respectively for the French and Moroccan clinics).

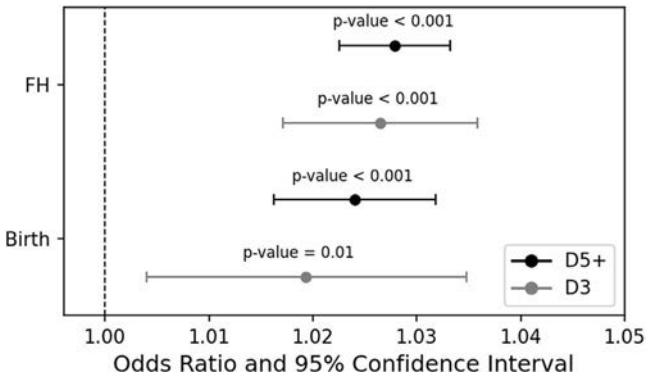


FIGURE 2 Odds ratio and 95% confidence interval resulting from the logistic regressions between the EMBRYOLY video score and transfer outcome across all centres per transfer day ($n = 2078$ blastocyst transfers and $n = 579$ cleavage transfers with known clinical pregnancy outcome; $n = 1646$ blastocyst transfers and $n = 475$ cleavage transfers with known live birth outcome). D3, cleavage stage transfer on day 3; D5+, blastocyst transfer on day 5, 6 or 7; FH, fetal heart.

EMBRYOLY's ranking could have statistically decreased the number of cycles required to reach a clinical pregnancy and increased the first cycle clinical pregnancy rate

When EMBRYOLY was considered on transferred and non-transferred embryos as a standalone, its highest ranked embryo matched the embryo transferred first by embryologists in 576 out of 926 egg retrievals (62.2% agreement rate). On average, these cohorts of embryos contained 5.37 ± 2.84 embryos in total, i.e. including non-transferred embryos; this agreement is, therefore, different from a random agreement (which would have been 18.6%). When EMBRYOLY did not agree with the embryo transferred first by the embryologist, the gap between the EMBRYOLY score of the embryo ranked highest by the embryologist and that ranked highest by EMBRYOLY was −0.14 ± 0.12 (the embryo selected by EMBRYOLY having, by definition, a necessarily higher score than the embryo selected by the embryologist).

When considering the use of EMBRYOLY as a support to embryologist on transferred blastocysts, their hypothetical FCP was 44.1% versus 19.8% for embryologists alone (McNemar's test: $P < 0.001$). This corresponds to 2.01 cycles to clinical pregnancy based on embryologists alone versus 1.66 cycles had the pre-selection of embryologists been also ranked by EMBRYOLY (Wilcoxon test: $P < 0.001$) (FIGURE 3). This was computed in cohorts of blastocysts with multiple single transfers of which at least one led to a clinical pregnancy and one that did not ($n = 260$ transferred embryos, from 111 egg retrievals).

DISCUSSION

To the best of our knowledge, this is the first study that evaluates an artificial intelligence-powered software for embryo evaluation as an adjunct tool to embryologists, and not as a standalone that would replace the embryologist's ranking entirely. Therefore, it is also the first study that estimates retrospectively the potential effect of an artificial intelligence-powered software on the number of cycles required to reach a clinical pregnancy in cohorts of sibling embryos when used in context of the embryologist's initial pre-selection. Several groups, including ours (*Duval* et al., *2023*), have shown that algorithms can help predict whether an embryo has a chance to lead to a clinical pregnancy or not across large datasets of embryos with areas under the curve ranging from 0.64 to 0.77 (*Kragh* et al., *2019*; *VerMilyea et al., 2020*; *Enatsu* et al., *2022*; *Liu* et al., *2023*; *Theilgaard Lassen* et al., *2023*). This, however, does not conclusively demonstrate that such algorithms can help embryologists rank sibling embryos. This is

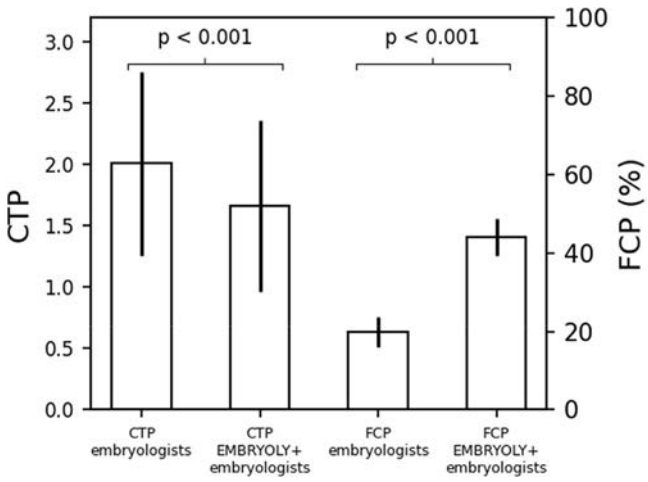


FIGURE 3 Comparison of cycles to clinical pregnancy (CTP) and first cycle clinical pregnancy rate (FCP) obtained by embryologists alone and EMBRYOLY plus embryologists on 260 blastocysts from 111 egg retrievals. *P*-values reported for CTP and FCP correspond respectively to a Wilcoxon and McNemar's test. Error bars represent the standard deviation for CTP and the standard error for FCP.

crucial because embryologists are not just expected to predict whether an embryo will lead to a pregnancy. Rather, they are expected to discern which embryo has the highest chance of leading to a pregnancy compared with the only other available sibling embryos. While area under the curve is supposed to illustrate ranking abilities, it does not conclusively demonstrate ranking abilities within restricted cohorts of sibling embryos, which represents the real conditions embryologists have to face when making a decision in the labroratory. Only one study (*Cimadomo* et al., *2023*) has investigated the ranking abilities of an artificial intelligence ranking tool in cohorts of sibling embryos, which had been biopsied for PGT-A. In addition, this study considered the artificial intelligence ranking tool as a standalone that would replace embryologists. Our study considers the effect of an artificial intelligence tool combined with the expertise of the embryologist showing that the model could have in fact helped embryologists better rank embryos in terms of their likelihood of clinical pregnancy and live birth with no prior knowledge of PGT-A results, which could have led to a higher first cycle clinical pregnancy rate and a shorter number of cycles required to lead to a clinical pregnancy.

The first attempt to estimate quantitatively the effect of an artificial intelligence tool in cohorts of embryos consisted of computing the time to clinical pregnancy of an artificial intelligence-powered algorithm trained on single images of blastocysts in fictional cohorts of embryos that were randomly assembled from a large dataset (*Diakiw* et al., *2022*). This evaluates the artificial intelligence model as if it were going to be followed blindly; it also only compared the putative time to clinical pregnancy with the one that would have been obtained randomly, and not with those resulting from the ranking of an embryologist. In addition, because they were not real sibling embryos, it assumed the embryos grouped randomly in a fictitious cohort would have had the same clinical pregnancy outcome when transferred to the same hypothetical patient, when, in reality, their transfer outcome heavily depended on the uterine and endometrial environment which was not controlled for from one patient to another. More recently, *Cimadomo* et al. *(2023)* have evaluated the ranking of iDAScore v1.0 across real sibling embryos and demonstrated that iDAScore v1.0 could have picked the euploid embryo that had led to a live birth earlier than embryologists in 15% of the cohorts that contained at least two euploid blastocysts, while prioritizing the embryo that had not led to a live birth in only 3% of such cohorts. The investigators, however, did not quantify the potential effect on the final live birth rates, because, in 29% of such cohorts, iDAScore v1.0 had ranked a non-transferred embryo higher than an embryo that led to a live birth, thus making it impossible to quantify the actual effect on time to live birth. This is presumably because iDAScore v1.0 was also evaluated as standalone software in the hopes it could fully automate the final decision of which embryo to transfer. In contrast, our experiments were designed, such that the effect of EMBRYOLY was considered in the context of the initial ranking of the embryologist, because it is reasonable to suppose that, in the short term, artificial intelligence models will be used as a complementary tool for embryologists, who will most likely still rely on their own clinical expertise and use artificial intelligence to refine their ranking (*Fitz* et al., *2021*). Also, in 64% of the cycles used by *Cimadomo* et al. *(2023)*, there was no possible effect of iDAScore v1.0 on time to live birth, as these cycles were only composed of transfers that had led to a live birth (either because there was only one embryo transferred and it led to a live birth, or because there were multiple single embryo transfers and each led to a live birth). In other words, iDAScore v1.0 could not have possibly ranked highest an embryo that did not lead to a live birth, so it was necessarily going to be as effective as the embryologist. This makes the overall task easier on iDAScore v1.0. In contrast, the data used to evaluate the effect of EMBRYOLY as an adjunct tool was on purpose only made of patients that had gone through at least one failed and one successful transfer, such that EMBRYOLY had in fact the opportunity to be 'wrong', thus illustrating more realistically the risks an embryologist takes in trusting an artificial intelligence model. Interestingly, the raw data from *Cimadomo* et al. *(2023)* could be exploited in our study and re-used to assess iDAScore v1.0 as a support to the embryologists (transferred embryos only) (TABLE 1), showing that time to live birth would have decreased from 1.99 to 1.59 ($P < 0.001$), similar to what was observed for EMBRYOLY on clinical pregnancy in our dataset. Conversely, when considering EMBRYOLY as a standalone in our dataset (including embryos with unknown outcome (TABLE 1 and Supplementary Figure 1), EMBRYOLY was also found to be as effective as embryologists in reaching clinical pregnancy within the same number of transfers in 65% of the cases (versus 52% for iDAScore v1.0), and result in earlier pregnancies in 22% of the cases (versus 15% for iDAScore v1.0). This comparison is, however, hampered by several factors: first, the differences in the two datasets, given that the maternal age in their study was higher (38.7 versus 34.1 in our study). Second, *Cimadomo* et al. *(2023)*

TABLE 1 DATASET AND MAIN RESULTS OF PRESENT STUDY COMPARED WITH THE STUDY BY *CIMADOMO* ET AL. *(2023)*.

Complete dataset	EMBRYOLY (present study)	iDAScore v1.0 (*Cimadomo* et al., *2023*)
Egg retrievals, *n*	2666	1232
Oocyte age (mean ± STD), years	34.1 ± 4.6	38.7 ± 3.4
Embryos, *n*	11988	3604
Biopsied for PGT-A, %	<2	100
Ploidy status	Not considered	Euploid only
Evaluation of artificial intelligence as an adjunct device; embryos with known outcome only from cohorts with at least one successful and one failed transfer		
Egg retrievals, *n*	111	71
Embryos with known transfer outcome, *n*	260	170
CTP/CTB without artificial intelligence	2.01	1.99
CTP/CTB with artificial intelligence	1.66	1.59
Evaluation of artificial intelligence as a standalone; embryos with and without known outcome from cohorts with at least two candidate embryos for transfer and one successful transfer		
Egg retrievals, *n*	162	202
Embryos with known transfer outcome, *n*	319	321
% treatments in which artificial intelligence was as effective as embryologists*	65	52
% of treatments with earlier CP/LB had AI been used**	22	15
% of treatments with later CP/LB had AI been used***	4	3
% of treatments with unknown effect of artificial intelligence	8	29

AI, artificial intelligence; CP/LB, clinical pregnancy/live birth; CTP/CTB, cycles to clinical pregnancy/live birth; PGT-A, preimplantation genetic testing for aneuploidy.

* AI and embryologist would have been equally effective since both would have obtained a successful transfer after the same number of transfers.

** AI would have identified a competent embryo earlier than the embryologist (that is, AI would have obtained a successful transfer with less transfers than the embryologist).

*** AI would have identified a competent embryo later than the embryologist (that is, AI would have obtained a successful transfer with more transfers than the embryologist).

measured time to live birth whereas time to clinical pregnancy was measured in the present study, because not all live birth results were known when the present study was conducted (Supplementary Table 1). Additionally, *Cimadomo* et al. *(2023)* only investigated the added value of iDAScore v1.0 when applied to euploid embryos, which presents two limitations. First, it limits the benefits of the selection tool to patients who have several euploid embryos. Second, by focusing solely on euploid embryos, the number of embryos to select from at the cohort level is reduced, as is the overall number of embryos with a known outcome. Consequently, eggs retrieved by *Cimadomo* et al. *(2023)* had fewer embryos eligible for transfer per cohort (1.4 ± 1.2 euploid blastocysts versus 6.8 ± 3.6 embryos for our data), and fewer transferred embryos per cohort (0.67 ± 0.79 versus 1.03 ± 1.01 on our data). Conversely, our study evaluated cohorts containing over 98% non-biopsied embryos, which may assist IVF centres in which PGT-A is not available or commonplace.

The results from this study suggest that EMBRYOLY, which was originally trained to predict clinical pregnancy, can help embryologists rank embryos in terms of their potential for clinical pregnancy but also live birth, even for embryos transferred at the cleavage stage. Ranking embryos on the third day of development can help clinics who do early transfer or help embryologists anticipate early on in their workflow which embryo might be transferred, frozen or later discarded. This study also illustrates the generalizability of the help EMBRYOLY can provide to embryologists given that results were statistically significant for data from two IVF centres that were never used in the training of the algorithm. This was the case for one centre from Morocco, which presented different demographics, notably a female population of younger age (Moroccan IVF centre 31.9 ± 4.7 years versus complete population 34.1 ± 4.6 years) and higher body mass index (Moroccan IVF centre 26.4 ± 5.1 versus complete population 24.1 ± 4.5). This suggests that EMBRYOLY can support embryologists in standardizing practices and reducing subjectivity in embryo evaluation (Supplementary Table 2). Variability, however, exists between different evaluation methods and also between and within operators for a given method (*Paternot* et al., *2011*; *Sundvall* et al., *2013*; *Cimadomo* et al., *2022*). In contrast, an algorithm will give a reproducible recommendation, which in addition can be trained more objectively than embryologists by only having been trained with objective clinical pregnancy outcomes, much like EMBROLY, and not with morphokinetic criteria that remain subjectively defined. Consequently, EMBRYOLY has the inherent potential to assist embryologists in enhancing the objectivity of their assessments and in harmonizing the practices across centres and countries.

The present study hinges on three main hypotheses: first, that embryologists transfer the embryo they deem the most viable first, which seems reasonable; second, that the transfer outcome would have been the same if the embryo had been transferred fresh or frozen and in a different order; this assumption relies on the fact that the maternal age and infertility diagnoses are controlled across transfer cycles for a given patient. Transfer outcome was also affected by uterine and endometrial environment, which may vary across transfer cycles. On average, the time elapsed between the first and last transfer for a given patient was less than a year (314 ± 399 days). In addition, it is assumed that there was no difference in endometrial preparation and luteal phase support protocols across transfers. The third assumption was that after their initial selection of embryos eligible for transfer, embryologists would have followed exactly the recommendation of EMBRYOLY in their decision making. While this will not always be the case, it is the most thorough approach to evaluating the risks an embryologist would take if they followed an artificial intelligence recommendation. Future studies should evaluate whether embryologists, after pre-selection of

embryos eligible for transfer, would have actually followed EMBRYOLY's recommendation prospectively to assess their readiness to trust such scores in their routine workflow. At any rate, a prospective study would make these conclusions stronger but would still only apply to transferred embryos. A limitation to always keep in mind revolves around the bias in data selection: here, IVF centres were asked to upload data of their choice and were encouraged to upload complete cohorts. On average, 4.5 embryos per cohort had been uploaded. This limitation could potentially affect the representativeness of this study: if an IVF centre only ever transfers good-quality embryos, only good-quality embryos will be evaluated for that centre. This risk is, however, mitigated by the multi-centric nature of the study, which adds diversity to the data sourcing. A final limitation of this study rests on the fact that EMBRYOLY has been compared to a decision made by one embryologist at a time, which does not consider the potential variability from one embryologist to another. At any rate, the embryologists involved ranged in seniority levels, from 5 to 20 years of experience, and more than 11 embryologists were involved, which is significantly diverse.

In conclusion, this study demonstrates how EMBRYOLY can help embryologists across multiple TLS, IVF centres and countries in the critical task of identifying the most promising embryos within a cohort in the hope of minimizing the time to clinical pregnancy. Much like any other ranking method, it aims to increase the first cycle clinical pregnancy rate, and not cumulative success rates. An improved ranking, however, could increase cumulative success rates by minimizing the number of patients who drop out after their first failed treatment and who could have become pregnant upon the transfer of another embryo from the cohort. Also, providing non- invasive and immediate ranking could promote fresh transfer for some patient populations (low and normal responders) in whom fresh transfer and the absence of PGT-A have been shown to lead to higher cumulative live birth rate (*Acharya* et al., *2018*).

Increasing the first cycle pregnancy rate is paramount to patients who want to become parents faster and to save time and minimize costs. Minimizing repeated failed transfers also reduces their physical and emotional strain. This strengthens the effect this tool can have in diverse and clinically relevant settings both on IVF centres to make their embryo evaluation less cumbersome and variable and on patients who can improve their chances of not going through repeated failed transfers.

DATA AVAILABILITY

The embryo videos and other patient data collected in this study are not publicly available owing to reasonable ethics and privacy concerns, and are not redistributable. The raw data for the final experiment is shared as supplementary table 3. For any interested collaborators, please contact the corresponding author. The AI model developed in this article is available for commercial use as part of ImVitro's software. The computer code developed is not publicly available owing to commercial restrictions.

AUTHORS' ROLES

ND and AB-C conceived the study and designed the methodology; AB-C was responsible for project management and supervision of research activity; ND was responsible for data curation, performing the research and formal analysis; ND and AB-C wrote the manuscript;

XPV, DN, BK, PS, CO, CG-S were involved in the data interpretation, and the review and editing of the final manuscript.

ACKNOWLEDGEMENTS

The authors want to thank Dr Emmanuel Chamorey for his help with our statistical analysis.

FUNDING

Funding for the study was provided by ImVitro, with grant funding received in part from BPIFrance Aide au Développement DeepTech (DOS0152872/00).

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103887.

REFERENCES

Acharya, K.S., Acharya, C.R., Bishop, K., Harris, B., Raburn, D., Muasher, S.J., 2018. Freezing of all embryos in in vitro fertilization is beneficial in high responders, but not intermediate and low responders: an analysis of 82,935 cycles from the Society for Assisted Reproductive Technology registry. Fertil. Steril. 110, 880–887. https://doi.org/10.1016/j.fertnstert.2018.05.024.

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Hum. Reprod. 26, 1270–1283. https://doi.org/10.1093/humrep/der037.

Barnes, J., Brendel, M., Gao, V.R., Rajendran, S., Kim, J., Li, Q., Malmsten, J.E., Sierra, J.T., Zisimopoulos, P., Sigaras, A., Khosravi, P., Meseguer, M., Zhan, Q., Rosenwaks, Z., Elemento, O., Zaninovic, N., Hajirasouliha, I., 2023. A non-invasive artificial intelligence approach for the prediction of human blastocyst ploidy: a retrospective model development and validation study. Lancet Digit Health 5, e28–e40. https://doi.org/10.1016/S2589-7500(22)00213-8.

Chavez-Badiola, A., Flores-Saiffe-Farías, A., Mendizabal-Ruiz, G., Drakeley, A.J., Cohen, J., 2020. Embryo Ranking Intelligent Classification Algorithm (ERICA): artificial intelligence clinical assistant predicting embryo ploidy and implantation. Reprod. Biomed. Online 41, 585–593. https://doi.org/10.1016/j.rbmo.2020.07.003.

Chen, T.-J., Zheng, W.-L., Liu, C.-H., Huang, I., Lai, H.-H., Liu, M., 2019. Using Deep Learning with Large Dataset of Microscope Images to Develop an Automated Embryo Grading System. FandR 01, 51–56. https://doi.org/10.1142/S2661318219500051.

Cimadomo, D., Chiappetta, V., Innocenti, F., Saturno, G., Taggi, M., Marconetto, A., Casciani, V., Albricci, L., Maggiulli, R., Coticchio, G., Ahlström, A., Berntsen, J., Larman, M., Borini, A., Vaiarelli, A., Ubaldi, F.M., Rienzi, L., 2023. Towards Automation in IVF: Pre-Clinical Validation of a Deep Learning-Based Embryo Grading System during PGT-A Cycles. J. Clin. Med. Res. 12. https://doi.org/10.3390/jcm12051806.

Cimadomo, D., Sosa Fernandez, L., Soscia, D., Fabozzi, G., Benini, F., Cesana, A., Dal Canto, M.B., Maggiulli, R., Muzzì, S., Scarica, C., Rienzi, L., De Santis, L., 2022. Inter-centre reliability in embryo grading across several IVF clinics is limited: implications for embryo selection. Reprod. Biomed. Online 44, 39–48. https://doi.org/10.1016/j.rbmo.2021.09.022.

Diakiw, S.M., Hall, J.M.M., VerMilyea, M., Lim, A.Y.X., Quangkananurug, W., Chanchamroen, S., Bankowski, B., Stones, R., Storr, A., Miller, A., Adaniya, G., van Tol, R., Hanson, R., Aizpurua, J., Giardini, L., Johnston, A., Van Nguyen, T., Dakka, M.A., Perugini, D., Perugini, M., 2022. An artificial intelligence model correlated with morphological and genetic features of blastocyst quality improves ranking of viable embryos. Reprod. Biomed. Online 45, 1105–1117. https://doi.org/10.1016/j.rbmo.2022.07.018.

Duval, A., Nogueira, D., Dissler, N., Maskani Filali, M., Delestro Matos, F., Chansel-Debordeaux, L., Ferrer-Buitrago, M., Ferrer, E., Antequera, V., Ruiz-Jorro, M., Papaxanthos, A., Ouchchane, H., Keppi, B., Prima, P.-Y., Regnier-Vigouroux, G.,

Trebesses, L., Geoffroy-Siraudin, C., Zaragoza, S., Scalici, E., Sanguinet, P., Cassagnard, N., Ozanon, C., De La Fuente, A., Gómez, E., Gervoise Boyer, M., Boyer, P., Ricciarelli, E., Pollet-Villard, X., Boussommier-Calleja, A., 2023. A hybrid artificial intelligence model leverages multi-centric clinical data to improve fetal heart rate pregnancy prediction across time-lapse systems. Hum. Reprod. 38, 596–608. https://doi.org/10.1093/humrep/dead023.

Enatsu, N., Miyatsuka, I., An, L.M., Inubushi, M., Enatsu, K., Otsuki, J., Iwasaki, T., Kokeguchi, S., Shiotani, M., 2022. A novel system based on artificial intelligence for predicting blastocyst viability and visualizing the explanation. Reprod. Med. Biol. 21, e12443. https://doi.org/10.1002/rmb2.12443.

Fitz, V.W., Kanakasabapathy, M.K., Thirumalaraju, P., Kandula, H., Ramirez, L.B., Boehnlein, L., Swain, J.E., Curchoe, C.L., James, K., Dimitriadis, I., Souter, I., Bormann, C.L., Shafiee, H., 2021. Should there be an "AI" in TEAM? Embryologists selection of high implantation potential embryos improves with the aid of an artificial intelligence algorithm. J. Assist. Reprod. Genet. 38, 2663–2670. https://doi.org/10.1007/s10815-021-02318-7.

Fukunaga, N., Sanami, S., Kitasaka, H., Tsuzuki, Y., Watanabe, H., Kida, Y., Takeda, S., Asada, Y., 2020. Development of an automated two pronuclei detection system on time-lapse embryo images using deep learning techniques. Reprod. Med. Biol. 19, 286–294. https://doi.org/10.1002/rmb2.12331.

Kan-Tor, Y., Ben-Meir, A., Buxboim, A., 2020. Can deep learning automatically predict fetal heart pregnancy with almost perfect accuracy? Hum. Reprod.. https://doi.org/10.1093/humrep/deaa083.

Khosravi, P., Kazemi, E., Zhan, Q., Malmsten, J.E., Toschi, M., Zisimopoulos, P., Sigaras, A., Lavery, S., Cooper, L.A.D., Hickman, C., Meseguer, M., Rosenwaks, Z., Elemento, O., Zaninovic, N., Hajirasouliha, I., 2019. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. npj Digital Medicine 2, 1–9. https://doi.org/10.1038/s41746-019-0096-y.

Kragh, M.F., Karstoft, H., 2021. Embryo selection with artificial intelligence: how to evaluate and compare methods? J. Assist. Reprod. Genet. 38, 1675–1689. https://doi.org/10.1007/s10815-021-02254-6.

Kragh, M.F., Rimestad, J., Berntsen, J., Karstoft, H., 2019. Automatic grading of human blastocysts from time-lapse imaging. Comput. Biol. Med. 115, 103494. https://doi.org/10.1016/j.compbiomed.2019.103494.

Liao, Q., Zhang, Q., Feng, X., Huang, H., Xu, H., Tian, B., Liu, J., Yu, Q., Guo, N., Liu, Q., Huang, B., Ma, D., Ai, J., Xu, S., Li, K., 2021. Development of deep learning algorithms for predicting blastocyst formation and quality by time-lapse monitoring. Commun Biol 4, 415. https://doi.org/10.1038/s42003-021-01937-1.

Li, K., Wang, Y., Gao, P., Song, G., Liu, Y., Li, H., Qiao, Y., 2022. UniFormer: Unified Transformer for Efficient Spatiotemporal Representation Learning. arXiv [cs.CV].

Liu, H., Zhang, Z., Gu, Y., Dai, C., Shan, G., Song, H., Li, D., Chen, W., Lin, G., Sun, Y., 2023. Development and evaluation of a live birth prediction model for evaluating human blastocysts from a retrospective study. Elife 12. https://doi.org/10.7554/eLife.83662.

Miyagi, Y., Habara, T., Hirata, R., Hayashi, N., 2020. Predicting a live birth by artificial intelligence incorporating both the blastocyst image and conventional embryo evaluation parameters. Artificial Intelligence in Medical Imaging 1, 94–107 . https://doi.org/10.35711/aimi.v1.i3.94.

Paternot, G., Wetzels, A.M., Thonon, F., Vansteenbrugge, A., Willemen, D., Devroe, J., Debrock, S., D'Hooghe, T.M., Spiessens, C., 2011. Intra- and interobserver analysis in the morphological assessment of early stage embryos during an IVF procedure: a multicentre study. Reprod. Biol. Endocrinol. 9, 127. https://doi.org/10.1186/1477-7827-9-127.

Petersen, B.M., Boel, M., Montag, M., Gardner, D.K., 2016. Development of a generally applicable morphokinetic algorithm capable of predicting the implantation potential of embryos transferred on Day 3. Hum. Reprod. 31, 2231–2244. https://doi.org/10.1093/humrep/dew188.

Sundvall, L., Ingerslev, H.J., Breth Knudsen, U., Kirkegaard, K., 2013. Inter- and intra-observer variability of time-lapse annotations. Hum. Reprod. 28, 3215–3221. https://doi.org/10.1093/humrep/det366.

Theilgaard Lassen, J., Fly Kragh, M., Rimestad, J., Nygård Johansen, M., Berntsen, J., 2023. Development and validation of deep learning based embryo selection across multiple days of transfer. Sci. Rep. 13, 4235. https://doi.org/10.1038/s41598-023-31136-3.

Tran, D., Cooke, S., Illingworth, P.J., Gardner, D.K., 2020. Reply: Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. Hum. Reprod.. https://doi.org/10.1093/humrep/dez264.

Tran, D., Cooke, S., Illingworth, P.J., Gardner, D.K., 2019. Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. Hum. Reprod. 34, 1011–1018. https://doi.org/10.1093/humrep/dez064.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need. arXiv [cs.CL].

VerMilyea, M., Hall, J.M.M., Diakiw, S.M., Johnston, A., Nguyen, T., Perugini, D., Miller, A., Picou, A., Murphy, A.P., Perugini, M., 2020. Development of an artificial intelligence-based assessment model for prediction of embryo viability using static images captured by optical light microscopy during IVF. Hum. Reprod. 35, 770–784. https://doi.org/10.1093/humrep/deaa013.

Zabari, N., Kan-Tor, Y., Or, Y., Shoham, Z., Shufaro, Y., Richter, D., Har-Vardi, I., Ben-Meir, A., Srebnik, N., Buxboim, A., 2023. Delineating the heterogeneity of embryo preimplantation development using automated and accurate morphokinetic annotation. J. Assist. Reprod. Genet. 40, 1391–1406. https://doi.org/10.1007/s10815-023-02806-y.

Received 7 September 2023; received in revised form 15 January 2024; accepted 5 February 2024.

RBMO

ELSEVIER

SHORT COMMUNICATION

Fast and furious: successful survival and resumption of meiosis in immature human oocytes vitrified and warmed using a short protocol

BIOGRAPHY

Juergen Liebermann, PhD, HCLD has been an IVF Laboratory Director since 1996. He received his doctoral degree from the Technical University Munich (Germany) in 1995, and his postdoctorate degree from the University of Wuerzburg (Germany) in 2004. His research focuses on optimizing egg and embryo vitrification to improve clinical outcomes.

J. Liebermann*, R. Brohammer, Y. Wagner, R. Smith, K. Even, J. Hirshfeld-Cytron, M.L. Uhler

ABSTRACT

Research question: Can immature oocytes vitrified and warmed using a short protocol survive and resume meiosis?

Design: This study examined modifications of oocyte vitrification and warming protocols that reduce the length of exposure to vitrification and warming solutions. In total, 561 germinal vesicles and 218 metaphase I oocytes that were immature at oocyte retrieval were vitrified at room temperature for 2 min. Warming was performed at 37°C for 2 min. Resumption of meiotic activity was evaluated after 24 and 48 h of culture. Two different commercially available vitrification and warming kits were used for comparison.

Results: Ninety-five percent of germinal vesicles survived, with no difference observed between the kits. The survival of metaphase I oocytes was, on average, 95.4% and did not differ significantly between the kits. Of the 533 germinal vesicles that survived, 491 converted to metaphase I oocytes (92.1%). After culture for 48 h, 54.4% converted to metaphase II oocytes. In addition, of the 208 metaphase I oocytes that survived warming, 84.1% converted to metaphase II oocytes after 24 h of culture. These maturation rates were similar to those of non-vitrified oocytes.

Conclusions: Vitrification and warming of oocytes at different nuclear maturation stages can be performed with 2 min of exposure to hypertonic solution and 2 min of exposure to hypotonic solution, respectively. This approach reduces exposure of the oocytes to room temperature during dehydration and rehydration. Warming in 0.5M sucrose helps to maintain and support the potential of oocytes to resume nuclear meiotic activity, and conversion from germinal vesicles to metaphase I and metaphase II oocytes.

INTRODUCTION

Vitrification and warming protocols for human oocytes are an integral part of IVF at many centres worldwide. However, the procedures are time-consuming, with multiple steps involved. A common practice is to start with six oocytes, expose them to equilibration and vitrification solutions for 16 min, and subsequently load two oocytes on to each carrier. Four oocytes are usually warmed at a time, with the process lasting 11 min. Under these circumstances, oocytes are exposed to room temperature during dehydration and rehydration, which can compromise survival rates and developmental potential post warming. However, recent studies have shown that oocytes can be vitrified in a much shorter

Fertility Centers of Illinois – US Fertility, Chicago, IL, USA

*Corresponding author. E-mail address: juergen.liebermann@fcionline.com (J. Liebermann). https://doi.org/10.1016/j.rbmo.2024.103976 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS
Oocytes
Vitrification
Short dehydration
Short rehydration
Efficiency

equilibration time (*Galardo et al., 2019*). Due to the rising demand for oocyte freezing, there is a need to make oocyte vitrification and warming protocols more consistent and efficient (*Brewer et al., 2023*; *Parmegiani et al., 2018*). Simplification and streamlining of warming protocols for human blastocysts have been demonstrated (*Liebermann et al., 2024*; *Taylor et al., 2022*).

MATERIALS AND METHODS

Fresh immature oocytes donated by patients to the IVF Laboratory for training and education purposes were used for this study, and consent was granted by each patient (ethical approval granted by Fertility Centers of Illinois, IRB ID 120160615, initial approval date 21 May 2022). In total, 561 germinal vesicles and 218 metaphase I oocytes that were immature at oocyte retrieval were cryopreserved using Vit Kit Freeze (FujiFilm Irvine Scientific, USA) and RapidVit-Oocyte (Vitrolife, Sweden). These two media were chosen because of their different compositions of equilibration solution and vitrification solution. However, 0.5M sucrose was used for both warming solutions.

Both vitrification media contain ethylene glycol, but the Irvine medium contains dimethyl sulfoxide (DMSO), while the Vitrolife medium contains propanediol (PDO). The Irvine equilibration solution is 7.5% ethylene glycol and 7.5% DMSO, and the vitrification solution is 15% ethylene glycol and 15% DMSO with 0.5M sucrose. The Vitrolife equilibration solution is 8% ethylene glycol and 8% PDO, and the vitrification solution is 16% ethylene glycol and 16% PDO with 0.55M sucrose. At room temperature, six oocytes were rinsed briefly in 100 μl of wash solution (modified human tubal fluid), exposed to a 200-μl drop of equilibration solution for 1 min, and then exposed to 2 × 100-μl drops of vitrification solution for 1 min. All six oocytes were loaded in one drop, deposited on a S-Cryolock (Biotech, USA), plunged into liquid nitrogen, and stored for 1–2 weeks until warming. Warming was performed at 37°C in a 500-μl drop of 0.5M sucrose (Dilution solution, Irvine Vit Kit Thaw; Warm 2, Vitrolife RapidWarm-Oocyte) for 1 min, followed by 1 min in 2 × 200-μl drops of wash solution (Wash solution, Irvine Vit Kit Thaw; Warm 4, Vitrolife RapidWarm-Oocyte). Post-warming survival and maturation after 24 and 48 h were evaluated by culturing in Global HighProtein media (LifeGlobal Group, USA). In addition, a control group of 65 non-vitrified germinal vesicles and 24 non-vitrified fresh metaphase I oocytes from retrievals at the time of warming, cultured for 24 and 48 h, was included to track and compare their maturation with the vitrified oocytes that survived.

Statistical analysis was carried out using chi-squared tests. Results are expressed as mean and SD, or as a percentage. Statistical significance was defined as $P < 0.05$.

RESULTS

No significant difference in the survival of germinal vesicles was found between the Irvine (242/255, 94.9%) and Vitrolife (291/306, 95.0%; $P = 0.93$) media (TABLE 1). The survival of metaphase I oocytes was also comparable between the two kits (Irvine: 112/117, 95.7%; Vitrolife: 96/101, 95.0%; $P = 0.81$). Due to the similar oocyte survival rates, the oocytes in both groups were combined to examine the resumption of meiotic activity and conversion rates. Of the 561 germinal vesicles, 533 (95.0%) survived, and 491 (92.1%) converted to metaphase I oocytes after 24 h of culture. After culture for 48 h, 54.4% (290/533) converted to metaphase II oocytes. Combining the vitrified and warmed metaphase I oocytes from both groups, 95.4% (208/218) survived warming, and 175 (84.1%) converted to metaphase II oocytes after 24 h of culture (TABLE 1). In the control group of 65 non-vitrified germinal vesicles, 93.8% converted to metaphase I oocytes after 24 h of culture ($P = 0.97$), and 58.5% converted to metaphase II oocytes after 48 h of culture ($P = 0.56$), which was similar to the conversion of vitrified–warmed germinal vesicles. In the control group of 24 non-vitrified metaphase I oocytes, 87.5% converted to metaphase II oocytes after 24 h of culture ($P = 0.94$).

DISCUSSION

Vitrification and warming of oocytes at different developmental stages can be achieved with 2 min of exposure to a hypertonic solution and 2 min of exposure to a hypotonic solution, respectively. This approach reduces exposure of the oocytes to room temperature during dehydration and rehydration, and survival rates are consistently >90%. The protocol helps to maintain the meiotic activity and maturation potential of oocytes by demonstrating high conversion rates from

TABLE 1 (A) SURVIVAL OF HUMAN OOCYTES AFTER 2 MIN OF VITRIFICATION AND 2 MIN OF WARMING IN TWO DIFFERENT MEDIA, AND (B) RESUMPTION OF MEIOTIC ACTIVITY OF VITRIFIED GERMINAL VESICLES AND METAPHASE I OOCYTES 24 H AND 48 H POST WARMING

(A) Survival	GV		*P*-value	MI oocyte		*P*-value
	Irvine	Vitrolife		Irvine	Vitrolife	
n	255	306		117	101	
Survival	242	291	0.93	112	96	0.81
%	94.9	95.0		95.7	95.0	

(B) Culture	Total GV survived	MI oocyte conversion 24 h later	MII oocyte conversion 48 h later	Total MI oocytes survived	MII oocyte conversion 24 h later
n	533	491	290	208	175
%	95.0	92.1	54.4	95.4	84.1

Survival data compared by chi-squared test.

GV, germinal vesicles; MI, metaphase I; MII, metaphase II.

metaphase I to metaphase II oocytes. The ability to load a maximum of six oocytes per carrier without compromising survival, combined with the decrease in time needed to perform vitrification and warming of oocytes, resulted in increased efficiency. However, it is recommended that four oocytes should be loaded per carrier in order to minimize limitations and risk, which increase with loading so many oocytes on one device. Warming in 0.5M sucrose medium does not compromise survival, and allows resumption of meiotic activity. However, increasing the speed of the procedure may have an impact on safety. Potential long-term side effects and drawbacks of minimizing the equilibration and dilution steps need careful follow-up. With the steady increase in the number of patients requiring oocyte cryopreservation in IVF clinics worldwide for medical and non-medical fertility preservation, embryology laboratories need more efficient and consistent oocyte vitrification and warming protocols. Follow-up studies are needed using this protocol on mature vitrified-warmed oocytes, and to evaluate fertilization, embryonic development and clinical outcome.

DATA AVAILABILITY

No data was used for the research described in the article.

REFERENCES

Brewer, A., Guerrero, C., VerMilyea, L., 2023. Warming vitrified oocytes in a fraction of the current required time results in superior survival rates. The 12th Congress of the Asia Pacific Initiative on Reproduction (Aspir), Adelaide, Australia. Oral presentation #386.

Galardo, M., Saenz, J., Risco, R., 2019. Human oocytes and zygotes are ready for ultra-fast vitrification after 2 minutes of exposure to standard CPA solutions. Scientific Reports 9, 15986.

Liebermann, J., Hrvojevic, K., Hirshfeld- Cytron, J., Brohammer, R., Wagner, Y., Susralski, A., Jasulaitis, S., Chan, S., Takhsh, E., Uhler, ML., 2024. Fast and furious: pregnancy outcome with one-step rehydration in the warming protocol for human blastocysts. RBMOnline 48 (4), 1–8.

Parmegiani, L., Beilby, K.H., Arnone, A., Bernardi, S., Maccarini, A.M., Nardi, E., Cognigni, G.E., Filicori, M., 2018. Testing the efficacy and efficiency of a single "universal warming protocol" for vitrified human embryos: prospective randomized controlled trial and retrospective longitudinal cohort study. J Assist Reprod Genet 35 (10), 1887–1895.

Taylor, T.H., Manns, J., Katz, I., Patrick, J., Whelan, J., Katz, S., 2022. Ultrafast warming protocol demonstrates similar outcomes and significantly decreases embryology workload compared to standard warming protocols, a randomized control trial with euploid blastocysts. ASRM Scientific Congress Expo; Oct 22-26, Anaheim, California, USA. Fertil Steril 118 (4), P-96, Supplement E150.

Received 20 December 2023; received in revised form 20 March 2024; accepted 20 March 2024.

RBMO

ARTICLE

Modified natural cycle allows a window of 7 days for frozen embryo transfer planning

BIOGRAPHY

Carlos Alonso is a Reproductive Medicine Specialist at IVIRMA, Spain. He obtained his medical degree from the University of Barcelona, completed his obstetrics and gynaecology residency at Hospital General Universitario Gregorio Marañón and pursued a Master's Degree in Reproductive Medicine at Complutense University. His primary areas of interest encompass gynaecological endocrinology and reproductive surgery.

Carlos Alonso-Mayo[1,*], Graciela Kohls[1], Samuel Santos-Ribeiro[2], Sergio Reis Soares[2], Juan A. Garcia-Velasco[1,3,4]

KEY MESSAGE

Comparable clinical outcomes are achieved when frozen embryo transfer (FET) is planned with rHCG when the dominant follicle size ranges from 13 to 22 mm in a modified natural cycle. This allows a flexibility that may help both patients and fertility units when monitoring and scheduling for FET.

ABSTRACT

Research question: Should ovulation be triggered in a modified natural cycle (mNC) with recombinant human chorionic gonadotrophin (rHCG) as soon as a mean follicle diameter of 17 mm is visible, or is more flexible planning possible?

Design: This multicentre, retrospective, observational study of 3087 single frozen blastocyst transfers in mNC was carried out between January 2020 and September 2022. The inclusion criteria included endometrial thickness ≥7 mm and serum progesterone <1.5 ng/ml. The main outcome was ongoing pregnancy rate. Secondary end-points were pregnancy rate, implantation rate, clinical pregnancy rate and miscarriage rate. The mean follicle size at triggering was stratified into three groups (13.0–15.9, 16.0–18.9 and 19.0–22 mm).

Results: The baseline characteristics between the groups did not vary significantly for age, body mass index and the donor's age for egg donation. No differences were found in pregnancy rate (64.5%, 60.2% and 57.4%; $P = 0.19$), clinical pregnancy rate (60.5%, 52.8% and 50.6%; $P = 0.10$), implantation rate (62.10%, 52.9% and 51.0%; $P = 0.05$) or miscarriage rate (15.0%, 22.2%; and 25.0%; $P = 0.11$). Although ongoing pregnancy rate (54.9%, 46.8% and 43.1%; $P = 0.02$) varied significantly in the univariable analysis, it was no longer significant after adjustment for the use of preimplantation genetic testing for aneuploidies and egg donation.

Conclusions: The findings showed rHCG could be flexibly administered with a mean follicle size between 13 and 22 mm as long as adequate endometrial characteristics are met, and serum progesterone is <1.5 ng/ml. Considering the follicular growth rate of 1–1.5 mm/day, this approach could allow a flexibility for FET scheduling of 6–7 days, simplifying mNC FET planning in clinical practice.

[1] IVIRMA Global Research Alliance, IVIRMA Madrid, Madrid, Spain
[2] IVIRMA Global Research Alliance, IVIRMA Lisbon, Lisbon, Portugal
[3] IVIRMA Global Research Alliance, IVI Foundation, Instituto de Investigación Sanitaria La Fe, Valencia, Spain
[4] Rey Juan Carlos University, Madrid Spain
DOI of original article: http://dx.doi.org/10.1016/j.rbmo.2023.103775

*Corresponding author. E-mail address: Carlos.Alonso@ivirma.com (C. Alonso-Mayo). https://doi.org/10.1016/j.rbmo.2023.103774 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEYWORDS
Cycle scheduling
Endometrial preparation
Frozen embryo transfer
Modified natural cycle
Ongoing pregnancy rate

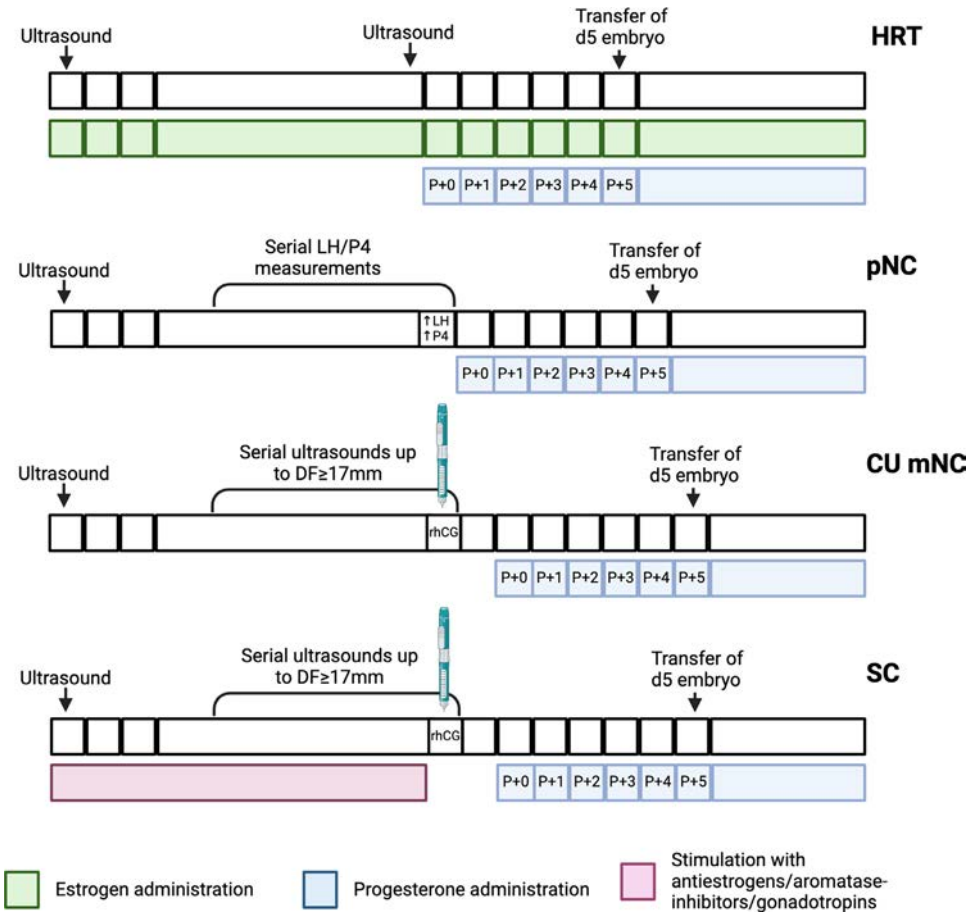


FIGURE 1 Different protocols used for frozen embryo transfer endometrial preparation. CU mNC, common use modified natural cycle; DF, dominant follicle; HRT, hormone replacement therapy;P4, progesterone; pNC, pure natural cycle; rhCG, recombinant human chorionic gonadotrophin; SC, stimulated cycle. Created with BioRender.com.

INTRODUCTION

Frozen embryo transfer (FET) cycles have been increasing globally over the last few years due to the advances in vitrification (*Cobo et al., 2021*; *De Geyter et al., 2020*; *Mackens et al., 2017*). The main contributors to this increase are cycle segmentation with the use of a gonadotrophin-releasing hormone agonist trigger to avoid ovarian hyperstimulation syndrome, the increase in use of preimplantation genetic testing for aneuploidies (PGT-A) and elective single-embryo transfer policies. Deferring embryo transfer may also enhance patient autonomy, allowing patients to better align this procedure with their work schedule, family or travel circumstances. However, the rise in FET cycles has also challenged reproductive units to seek optimal strategies for scheduling their workload when planning embryo thawing and transfer.

The two main methods for endometrial preparation for FET are hormone replacement therapy (HRT) and natural cycles (FIGURE 1) (*Mumusoglu et al., 2021*). No protocol has been demonstrated to be superior in terms of achieving pregnancy (*Ghobara et al., 2017*; *Glujovsky et al., 2020*; *Wolfe et al., 2023*). Traditionally, HRT has been the most common protocol for FET as it overcomes the variability found in menstrual cycles, requiring fewer controls and allowing flexibility at the time of embryo transfer (*Mackens et al., 2017*). Nevertheless, in this protocol patients must take oestrogen therapy, which may cause side effects, leading to reluctance to continue with it. Furthermore, a higher incidence of hypertensive disorders of pregnancy after HRT (*Ginström Ernstad et al., 2019*; *Saito et al., 2019*; *von Versen-Höynck and Griesinger, 2022*; *von Versen-Höynck et al., 2019*), an increased likelihood of large for gestational age or macrosomy (*Elias et al., 2020*; *Ginström Ernstad et al., 2019*; *Litzky et al., 2018*) and a higher incidence of postpartum haemorrhage and placenta-associated complications (*Busnelli et al., 2022*) have recently also been noted. Thus, although HRT may have been a convenient option in the past, natural cycle FET is progressively gaining interest in reproductive care units and has been posited by some as the preferred method for endometrial preparation for FET (*Lawrenz et al., 2020*).

Natural cycles can only be offered if ovarian function is present. In pure natural cycles (pNC) there is no pharmacological intervention; a serum or urinary LH surge and/or serum progesterone is detected, and the subsequent transfer is then scheduled accordingly (*Lawrenz et al., 2023*). Home monitoring by urine ovulation detection kits has been observed to be an effective and patient-friendly approach (*Zaat et al., 2023*); however, it entails a rigid scheme and does not allow flexible planning of FET. Furthermore, the temporal association between LH and progesterone rise has been observed to vary and it is not clear whether FET should be scheduled according to serum progesterone concentrations or LH

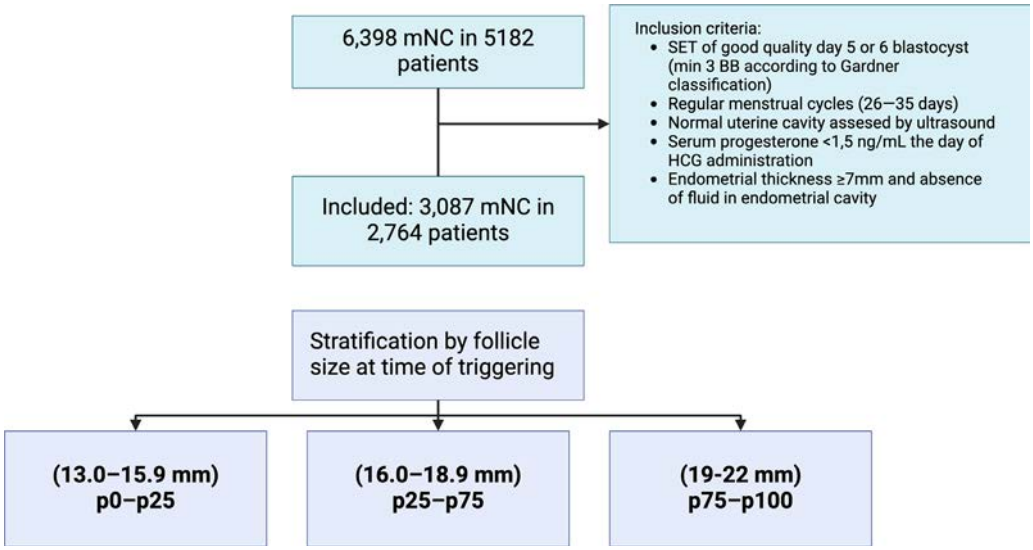


FIGURE 2 Patients and stratification. HCG, recombinant human chorionic gonadotrophin; mNC, modified natural cycle; SET, single-embryo transfer.

concentrations in a natural cycle (*Coughlan et al., 2023*).

Contrary to pNC, in modified natural cycles (mNC), final oocyte maturation is usually triggered with urinary or recombinant human chorionic gonadotrophin (rHCG), allowing a patient-friendly approach as it requires less monitoring than pNC. The standard mean diameter necessary for triggering in an mNC has classically been established as 17 mm, mimicking cycles aiming to obtain a mature oocyte in IVF ovarian stimulation. However, it is known that a mature oocyte is not needed for FET; what is necessary is a competent corpus luteum that produces oestrogens, progesterone and perhaps other substances such as vasoactive products like relaxin and vascular endothelial growth factor, among others (*Pereira et al., 2021*). Interestingly, however, research on triggering at different follicle sizes in an mNC has been limited.

The purpose of this study was to evaluate whether an mNC allows flexibility regarding follicle size when triggering with rHCG without hindering early pregnancy outcomes.

MATERIALS AND METHODS

A multicentre, retrospective, observational study of single frozen blastocyst transfers in mNC was carried out at IVIRMA centers in Spain, Portugal and Italy from January 2020 to September 2022 in women undergoing IVF using either autologous or donated oocytes.

The selection criteria were women who underwent a single frozen blastocyst transfer of a good-quality day 5 or 6 blastocyst (minimum 3BB according to the Gardner and Schoolcraft classification; *Gardner et al., 2000*) with regular menstrual cycles (from 26 to 35 days), a normal uterine cavity assessed by ultrasonography and the presence of all the following criteria on the day of rHCG administration: serum progesterone <1.5 ng/ml, endometrial thickness ≥7 mm and absence of fluid in the endometrial cavity. Cases in which the follicle size on the day of triggering with rHCG was not recorded were excluded from the analysis (FIGURE 2).

Patients were monitored by transvaginal ultrasonography starting on the first 3 days of menstruation to confirm ovarian quiescence. They were followed by transvaginal ultrasonography according to the physician's discretion, and progesterone and oestradiol concentrations were determined on the day of the last transvaginal ultrasound examination and triggering with rHCG. Administration of the trigger medication was decided according to physician's preference; it consisted of a single dose of subcutaneous rHCG (Ovitrelle 250 μg). Natural micronized progesterone 200 mg twice daily (Progeffik 200 mg or Utrogestan 200 mg) was started 2 days after the trigger according to the centres' standard protocol, and SET was performed on day rHCG+7.

The primary end-point was ongoing pregnancy rate (OPR). Secondary end-points were the pregnancy rate, implantation rate, clinical pregnancy rate (CPR) and miscarriage rate. The OPR was defined as each pregnancy showing a heartbeat on ultrasonography after 11 weeks of gestation divided by the number of transfers performed.

The pregnancy rate was calculated as the number positive serum beta-human chorionic gonadotrophin (β-HCG) tests divided by the number of transfers performed. Serum (β-HCG was considered positive if it was above 10 mIU/ml measured from the 10th day after embryo transfer.

The implantation rate was calculated as the number of gestational sacs diagnosed on ultrasonography divided by the number of transfers performed.

The CPR was calculated as the number of pregnancies diagnosed by the ultrasonographic visualization of one or more gestational sacs divided by the number of embryo transfers performed. The first ultrasound scan was predominantly performed during the 5th gestational week, and the differences between the implantation rate and CPR are explained by monozygotic twins.

The miscarriage rate was calculated considering the number of biochemical miscarriages (cases with a positive pregnancy test that did not show a gestational sac) and clinical miscarriages

(miscarriages happening since the detection of a gestational sac up to the 20th week of pregnancy) divided by the number of positive pregnancies.

Follicle size at time of triggering was stratified into three groups to make comparisons with the 'standard' follicle size. The stratification of follicle sizes comprised the smallest follicle quartile (p0–p25), the middle quartiles (p25–p75) and the largest follicle quartile (p75–p100). Comparisons were made between these three groups.

Additionally, the impact of follicle size on OPR was evaluated for each follicle size, and a receiver operating characteristic (ROC) curve was built in order to assess whether a particular follicle diameter had the best predictive value for OPR.

The Statistical Package for Social Sciences, version 20.0 (SPSS; IBM Corporation, USA) was used for the statistical analysis, and differences were considered significant if the probability of their occurrence by chance was less than 0.05.

Given the retrospective nature of this study, no formal sample size calculation could be carried out before the start of the study, so a post-hoc power calculation was made in relation to the primary end-point.

Continuous variables were expressed as mean values ± standard deviations (SD), while categorical variables were expressed as proportions (percentages), including 95% confidence intervals (CI). Baseline characteristics and main and secondary outcomes were compared using appropriate statistical tests (chi-squared test and analysis of variance). For OPR an adjustment in terms of the different distribution of basal characteristics in the study groups (use of PGT-A and egg donation) was carried out using an analysis of covariance test. Hormonal parameters were not taken into consideration as both oestradiol concentrations and progesterone concentrations were considered to depend on follicle size at the time of triggering.

The necessary information was exported from the electronic medical records, and the exported data were duly encrypted to protect patients' clinical and personal information as provided for by the applicable law (Law 14/2007 on Biomedical Research). The study was approved by the Institutional Review Board (2210-MAD-140-JG) on 7 November 2022.

RESULTS

From January 2020 until September 2022, 6398 mNC for FET were carried out in 5182 patients. After applying the inclusion and exclusion criteria, the final analysis included 2764 patients who underwent a total of 3087 FET in an mNC (FIGURE 2).

Follicle size at time of triggering was stratified into three groups (13.0–15.9 mm, 16.0–18.9 mm and ≥19.0 mm). The baseline characteristics of the study population are shown in TABLE 1. There were no significant differences in the women's age at the time of transfer, the body mass index or the age of the donor at the time of egg retrieval if egg donation had been used.

Regarding the early pregnancy results, no differences were seen in pregnancy rate, CPR, implantation rate or miscarriage rate, but differences were found in the OPR. OPR was found to be significantly different between the three groups (13–15.9 mm group, 54.8%; 16–18.9 mm group, 46.8%; 19.0–22 mm group, 43.1%; $P = 0.02$) (TABLE 2). However, after adjusting for the indication for PGT-A and oocyte donation, these differences were not maintained, as seen in TABLE 3.

Finally, OPR was assessed by follicle size (FIGURE 3) showing a similar OPR across the different diameters ($P = 0.07$). As shown in FIGURE 4, the ROC analysis did not show any predictive value of follicle size at the time of triggering for the OPR, with an area under the curve of 0.51 (CI 95% 0.47–0.55; $P = 0.56$).

The post-hoc statistical power for the primary end-point of the study was found to be 0.95.

DISCUSSION

This study demonstrated that, in an mNC, triggering with rHCG can be done at different follicular sizes – from 13 to 22 mm – as long as adequate endometrial characteristics are met and elevation of serum progesterone has not occurred, without any evident impact on the OPR. Furthermore, neither of the study's secondary end-points showed statistically significant differences, giving a window of flexible planning for women undergoing mNC.

TABLE 1 BASELINE CHARACTERISTICS OF THE STUDY GROUP FORZEN EMBRYO TRANSFER CYCLES

Characteristics	13.0–15.9 mm	16.0–18.9 mm	19.0–22 mm	All	*P*-value
n	124	2014	949	3087	–
Age (years)	38.7 ± 4.5	38.3 ± 4.3	38.4 ± 4.2	38.3 ± 4.3	0.4
Body mass index (kg/m^2)	23.2 ± 3.7	23.1 ± 3.8	23.3 ± 4.0	23.1 ± 3.9	0.6
Egg donation	39.5% (49/124)	27.9% (561/2014)	27.4% (260/949)	28.4% (876/3087)	0.02
PGT-A	19.4% (24/124)	34.0% (685/2014)	37.3% (354/949)	34.4% (1063/3087)	<0.001
Age of the donor (years)[a]	25.3 ± 2.1	25.1 ± 2.0	25.2 ± 2.3	25.1 ± 2.1	0.6
Cycle duration until rHCG (days)	10.2 ± 3.1	11.9 ± 3.2	12.4 ± 3.6	11.9 ± 3.2	0.01
Serum oestradiol on the day of rHCG (pg/ml)	219.8 ± 212.0	247.5 ± 113.5	323.0 ± 153.5	270.0 ± 136.6	<0.001
Serum progesterone on the day of rHCG (ng/ml)	0.33 ± 0.27	0.29 ± 0.27	0.35 ± 0.31	0.31 ± 0.28	<0.001

Data presented as mean ± standard deviation or percentage (*n*/*N*). Data for size groupings were compared using analysis of variance and chi-squared tests.

[a] If egg donation.

PGT-A, preimplantation genetic testing for aneuploidies; rHCG, recombinant human chorionic gonadotrophin.

TABLE 2 EARLY PREGNANCY RESULTS STRATIFIED BY FOLLICLE SIZE AT THE TIME OF TRIGGERING

Outcomes	13.0–15.9 mm	16.0–18.9 mm	19.0–22 mm	*P*-value
Pregnancy rate	64.5% (80/124)	60.2% (1212/2014)	57.4% (545/949)	0.2
Implantation rate	62.1% (77/124)	53.1% (1086/2014)	51.0% (484/949)	0.05
Clinical pregnancy rate	60.5% (75/124)	52.8% (1064/2014)	50.6% (480/949)	0.1
Miscarriage (biochemical and clinical)	15.0% (12/80)	22.2% (269/1212)	25.0% (136/545)	0.11
Biochemical miscarriage	6.3% (5/80)	12.2% (148/1212)	11.9% (65/545)	0.28
Clinical miscarriage	8.8% (7/80)	10.0% (121/1212)	13.0% (71/545)	0.14
Ongoing pregnancy rate	54.8% (68/124)	46.8% (943/2014)	43.1% (409/949)	0.02

Data are percentage (*n*/*N*). Data were compared using analysis of variance and chi-squared tests.

Pregnancy, implantation, clinical pregnancy and ongoing pregnancy rates were calculated using the number of embryo transfers performed. All miscarriage rates were calculated using the number of positive pregnancies. Clinical miscarriages were defined as a loss up to the 20th week of pregnancy calculated using the number of positive pregnancies. Ongoing pregnancy was defined as a pregnancy with heartbeat at 11 weeks.

TABLE 3 AOR FOR ONGOING PREGNANCY RATE AFTER ADJUSTING FOR POSSIBLE CONFOUNDERS

Follicle grouping comparisons	aOR	95% CI	*P*-value
13–15.9 mm (*n* = 124) vs 16–18.9 mm (*n* = 2014)	2.4	0.7–7.6	0.15
19–22 mm (*n* = 949) vs 16–18.9 mm (*n* = 2014)	0.8	0.5–1.1	0.1

Confounders controlled for were the use of preimplantation genetic testing for aneuploidies and egg donation.

aOR, adjusted odds ratio.

Whether endometrial thickness independently affects early pregnancy and obstetric results remains a subject of ongoing debate. There is no definitive consensus regarding the necessity for a cut-off of ≥7 mm for embryo transfer. In a recent retrospective study conducted by Ata and colleagues that encompassed 959 single-euploid FET, there was no decrease of live birth rate until the endometrial thickness was 4 mm (*Ata et al., 2023*). Nevertheless, it is worth noting that the prevailing practice currently favours the adoption of the 7 mm cut-off, which is why it was considered as an exclusion criterion in the current study.

Although many centres have a 7-day work schedule, meaning that cycle monitoring and embryo transfer can be carried out on any given day, it is preferable to arrange visits and procedures for weekdays, when a full complement of staff is available. To this extent, HRT endometrial preparation for FET currently presents some advantages over an mNC as it requires less monitoring and provides more flexibility when programming FET. This allows the moment of the embryo transfer to be better adapted to the clinic workload, besides

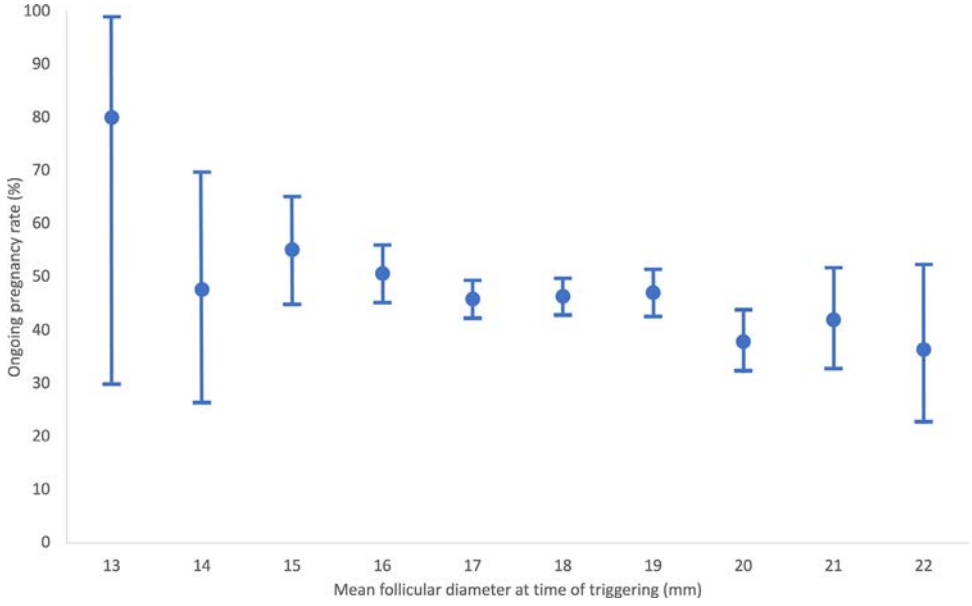


FIGURE 3 Ongoing pregnancy rate (OPR) by mean follicular diameter at the time of triggering. The data are shown as mean ± confidence interval 2.5–97.5%. The OPR was calculated as the number of pregnancies with a heartbeat at 11 weeks in relation to the total number of embryo transfers. Analysis of variance test ($P = 0.07$).

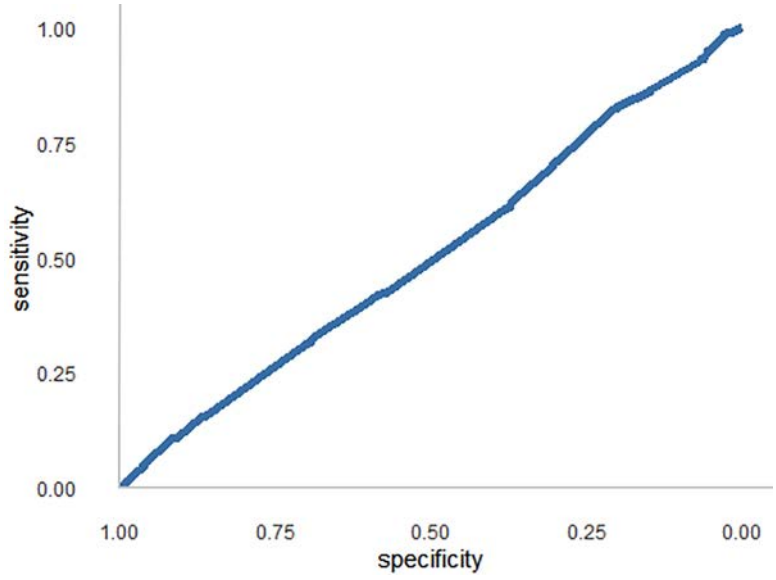


FIGURE 4 Receiver operating curve evaluating the effect of follicular size on ongoing pregnancy rate (OPR). Follicle size did not show a predictive value for OPR, with an area under the curve of 0.51 (CI 95% 0.47–0.55; $P = 0.56$).

enabling patients to better plan their travel or work leave if needed.

However, HRT endometrial preparation requires more medication than an mNC, including replacement doses of both oestradiol and progesterone, which potentially increases the costs and the risk of adverse effects. Regarding the latter, the higher rate of obstetric complications associated with HRT FET is thought to be due to the absence of a corpus luteum in HRT cycles (*Ginström Ernstad et al., 2019*; *Saito et al., 2019*; *von Versen-Höynck and Griesinger, 2022*; *von Versen-Höynck et al., 2019*). Specifically, the lack of a corpus luteum seems to impair maternal circulatory adaptation to pregnancy (*von Versen-Höynck and Griesinger, 2022*). Before the placenta has become established, the corpus luteum serves as a significant source of oestrogen, progesterone and vasoactive factors such as relaxin among others, which are postulated to be crucial for initial placentation (*Pereira et al., 2021*). Little is known about relaxin and other factors secreted by the corpus luteum into the circulation, but they may be important to guarantee correct placentation and cannot be replaced in artificial FET cycles with an absent corpus luteum.

In an mNC, on the other hand, a corpus luteum is present. Extrapolating from ovarian stimulation cycles, an appropriate follicle size at the time of triggering in an mNC has classically been established at 17 mm as this is the follicular size at which a mature oocyte is expected. However, for an FET cycle, the presence of a mature oocyte is not required. The only requirement is to have adequate endometrial development achieved by endogenous oestrogens, later progesterone exposure and/or any other factors beyond hormones secreted by the corpus luteum. This consideration led the current authors to contemplate the possibility that, once the endometrium is ready for embryo transfer, triggering with rHCG and FET planning could be carried out at different follicle sizes, if progesterone concentrations have not risen.

To the authors' knowledge, only two studies have evaluated this issue. Weiss and colleagues (*Weiss et al., 2021*) carried out an mNC protocol in 42 women in which progesterone administration was started solely in relation to endometrial characteristics as assessed by ultrasonography (endometrial thickness ≥7 mm and a trilaminar appearance) and the presence of a dominant follicle of at least 12 mm. Final oocyte maturation was not performed, the study included a very limited sample size, and the embryos were transferred at the cleavage stage, but an OPR of 54.7% (95% CI 39.1–70.5%) was reported. Godinho and co-workers (*Godinho et al., 2021*) performed a similar study observing even higher live birth rates in comparison with the conventional mNC protocol (32.7% versus 46.2%; $P = 0.02$); they included 79 patients in what they called the progesterone-programmed FET group, and 1155 women in the conventional mNC protocol group.

Although the current study did not include women with irregular cycles, an option for women with irregular or anovulatory cycles could be a mildly stimulated cycle using an oral agent (clomiphene citrate or letrozole) or exogenous gonadotrophins. Embryo transfer planning would then proceed as in the mNC protocol (*Mumusoglu et al., 2021*).

LH/chorionic gonadotrophin receptor (LHCGR) signalling plays an essential role in natural cycles through a transduction of the signal of LH or HCG. LHCGR is constitutively expressed on theca cells, while its expression in granulosa cells is induced by FSH. This expression can be observed from a follicular diameter as small as 3 mm, with higher expression noted when the follicular diameter exceeds 10 mm (*Jeppesen et al., 2012*; *Kishi et al., 2018*). LHCGR expression at these stages could be enough for the follicle to be able to ovulate, ultimately leading to the production of progesterone and other necessary substances by the corpus luteum.

This is the first study to evaluate early pregnancy results according to follicular size at the time of triggering in an mNC. The study's main strength is its sample size and the strict inclusion criteria, including only SET at the blastocyst stage.

The study has, however, certain limitations. Due to its observational retrospective design, it is possible that some confounding factors not included in the analysis may have affected the results. Furthermore, 84.7% of the mNCs included in this study involved a follicle size of 15–19 mm at the time of triggering, which leaves only a minority of cases in which triggering was carried out at 'non-conventional' follicle sizes. The decision of the physician to trigger with 'small' follicle sizes very often relied on the presence of a previous history of a cycle cancelled because of premature ovulation. This may confer different characteristics on this population compared with the 'standard' follicle size group. In fact, embryo transfers from egg donation cycles were more frequent in patients in whom rHCG administration occurred at follicle sizes

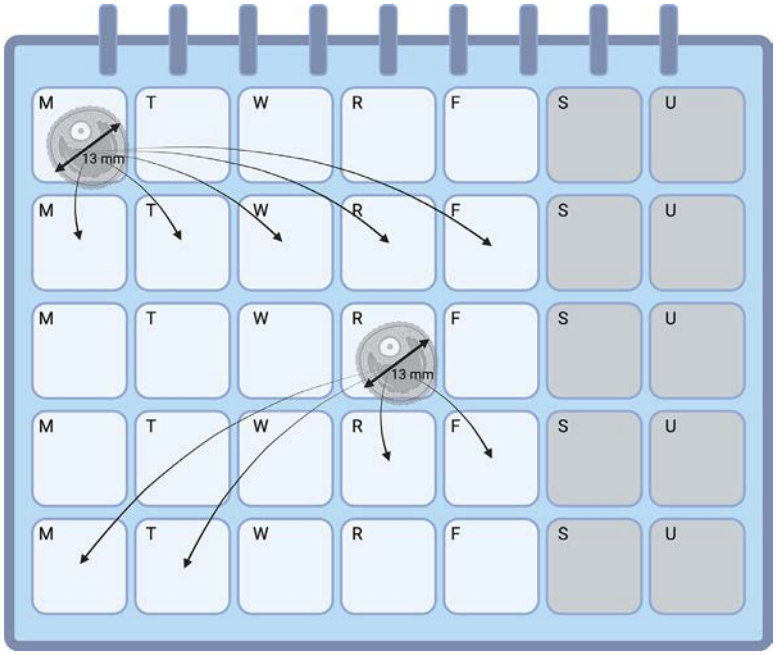


FIGURE 5 Schematic representation of the implications of flexibility in a modified natural cycle (mNC) for a blastocyst transfer. In this illustration two hypothetical cases are depicted once adequate endometrial characteristics on ultrasonography have been met. Considering that triggering with recombinant human chorionic gonadotrophin (rHCG) can be carried out once the dominant follicle reaches 13 mm in mean diameter, if this is achieved on a Monday, embryo transfer could be performed from the Monday until at least Friday of the following week. If a 13 mm diameter follicle is found on a Thursday, frozen embryo transfer could be scheduled from at least the next Thursday to the following Tuesday. This approach allows for simple embryo transfer scheduling evenly throughout the week, minimizing – or even avoiding – embryo transfers on weekends, and reducing the need for frequent follicle measurements. M, Monday; T, Tuesday; W, Wednesday; R, Thursday; F, Friday; S, Saturday; U, Sunday. Created with BioRender.com.

below 16 mm. These could have an impact on the results, as embryos derived from egg donation programmes are associated with higher implantation and OPR. In that sense, the 13–15.9 mm diameter group had the best OPR, although the statistical significance disappeared after adjustment for the use of PGT-A and the origin of the egg. Finally, the impact of follicular diameter on the risk of obstetric complications was not studied. This is a hypothesis-generating study; thus, these results need to be confirmed by future prospective studies.

Although LH was not considered a parameter in the study analysis, and despite its potential role in pregnancy chances, the authors believe it is unlikely to have affected the results. First, all the patients underwent evaluation for serum progesterone concentrations. In the event that LH concentrations were initiating a rise, it would probably have been very subtle, given that the progesterone concentration had to be below 1.5 ng/ml. Second, the pragmatic addition of vaginal progesterone 2 days after the trigger should have mitigated any differences in serum progesterone concentrations induced by the LH rise. Third, the initiation of the secretory transformation of the endometrium is promoted by progesterone exposure and not by LH concentrations. Further evaluation of LH concentrations in future studies may increase knowledge about the physiology in these early stages of pregnancy.

In conclusion, this study's results suggest that, once endometrial thickness reaches ≥7 mm and if progesterone is <1.5 ng/ml, FET in an mNC may be planned by triggering with rHCG when the size of the dominant follicle ranges from 13 mm to 22 mm, demonstrating a comparable clinical outcome across different follicular sizes. Considering a follicular growth rate of 1–1.5 mm per day (*Baerwald et al., 2009*), this approach would allow a flexible range of 5–7 days for FET planning in patients undergoing an mNC protocol. The results may help both patients and fertility units, simplifying cycle monitoring and facilitating FET scheduling on one hand, and minimizing the burden of multiple visits by patients on the other (FIGURE 5).

DATA AVAILABILITY

Data will be made available on request.

REFERENCES

Ata, B., Liñán, A., Kalafat, E., Ruíz, F., Melado, L., Bayram, A., Elkhatib, I., Lawrenz, B., Fatemi, H.M., 2023. Effect of the endometrial thickness on the live birth rate: insights from 959 single euploid frozen embryo transfers without a cutoff for thickness. Fertil. Steril. 120, 91–98. https://doi.org/10.1016/j.fertnstert.2023.02.035.

Baerwald, A.R., Walker, R.A., Pierson, R.A., 2009. Growth rates of ovarian follicles during natural menstrual cycles, oral contraception cycles, and ovarian stimulation cycles. Fertil. Steril. 91, 440–449. https://doi.org/10.1016/j.fertnstert.2007.11.054.

Busnelli, A., Schirripa, I., Fedele, F., Bulfoni, A., Levi-Setti, P.E., 2022. Obstetric and perinatal outcomes following programmed compared to natural frozen-thawed embryo transfer cycles: a systematic review and meta-analysis. Hum. Reprod. 37, 1619–1641. https://doi.org/10.1093/humrep/deac073.

Cobo, A., García-Velasco, J.A., Remohí, J., Pellicer, A., 2021. Oocyte vitrification for fertility preservation for both medical and nonmedical reasons. Fertil. Steril. 115, 1091–1101. https://doi.org/10.1016/j.fertnstert.2021.02.006.

Coughlan, C., Ata, B., Gallego, R.D., Lawrenz, B., Melado, L., Samir, S., Fatemi, H., 2023. Interindividual variation of progesterone elevation post LH rise: implications for natural cycle frozen embryo transfers in the individualized medicine era. Reprod. Biol. Endocrinol. 21, 47. https://doi.org/10.1186/s12958-023-01096-4.

De Geyter, C., Wyns, C., Calhaz-Jorge, C., de Mouzon, J., Ferraretti, A.P., Kupka, M., Nyboe Andersen, A., Nygren, K.G., Goossens, V., 2020. 20 years of the European IVF-monitoring Consortium registry: what have we learned? A comparison with registries from two other regions. Hum. Reprod. 35, 2832–2849. https://doi.org/10.1093/humrep/deaa250.

Elias, F.T.S., Weber-Adrian, D., Pudwell, J., Carter, J., Walker, M., Gaudet, L., Smith, G., Velez, M.P., 2020. Neonatal outcomes in singleton pregnancies conceived by fresh or frozen embryo transfer compared to spontaneous conceptions: a systematic review and meta-analysis. Arch. Gynecol. Obstet. 302, 31–45. https://doi.org/10.1007/s00404-020-05593-4.

Gardner, D.K., Lane, M., Stevens, J., Schlenker, T., Schoolcraft, W.B., 2000. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertil. Steril. 73, 1155–1158. https://doi.org/10.1016/S0015-0282(00)00518-5.

Ghobara, T., Gelbaya, T.A., Ayeleke, R.O., 2017. Cycle regimens for frozen-thawed embryo transfer. Cochrane Database Syst. Rev. 2017. https://doi.org/10.1002/14651858.CD003414.pub3.

Ginström Ernstad, E., Wennerholm, U.-B., Khatibi, A., Petzold, M., Bergh, C., 2019. Neonatal and maternal outcome after frozen embryo transfer: Increased risks in programmed cycles. Am. J. Obstet. Gynecol. 221,. https://doi.org/10.1016/j.ajog.2019.03.010 126.e1-126.e18.

Glujovsky, D., Pesce, R., Sueldo, C., Quinteiro Retamar, A.M., Hart, R.J., Ciapponi, A., 2020. Endometrial preparation for women undergoing embryo transfer with frozen embryos or embryos derived from donor oocytes. Cochrane Database Syst. Rev. 2020. https://doi.org/10.1002/14651858.CD006359.pub3.

Godinho, C., Soares, S.R., Santos-Ribeiro, S., 2021. Natural proliferative phase frozen embryo transfers (NPP-FET) – A new approach that facilitates scheduling without hindering pregnancy outcomes. Fertil. Steril. 116, e165. https://doi.org/10.1016/j.fertnstert.2021.07.455.

Jeppesen, J.V., Kristensen, S.G., Nielsen, M.E., Humaidan, P., Dal Canto, M., Fadini, R., Schmidt, K.T., Ernst, E., Yding Andersen, C., 2012. LH-Receptor Gene Expression in Human Granulosa and Cumulus Cells from Antral and Preovulatory Follicles. J. Clin. Endocrinol. Metab. 97, E1524–E1531. https://doi.org/10.1210/jc.2012-1427.

Kishi, H., Kitahara, Y., Imai, F., Nakao, K., Suwa, H., 2018. Expression of the gonadotropin receptors during follicular development. Reprod. Med. Biol. 17, 11–19. https://doi.org/10.1002/rmb2.12075.

Lawrenz, B., Ata, B., Kalafat, E., Melado, L., ElKhatib, I., Del Gallego, R., Fatemi, H., 2023. Are systemic progesterone levels in true natural cycle euploid frozen embryo transfers with luteal phase support predictive for ongoing pregnancy rates? Hum. Reprod. 38, 1318–1324. https://doi.org/10.1093/humrep/dead104.

Lawrenz, B., Coughlan, C., Melado, L., Fatemi, H.M., 2020. The ART of frozen embryo transfer: back to nature!. Gynecol. Endocrinol. 36, 479–483. https://doi.org/10.1080/09513590.2020.1740918.

Litzky, J.F., Boulet, S.L., Esfandiari, N., Zhang, Y., Kissin, D.M., Theiler, R.N., Marsit, C.J., 2018. Effect of frozen/thawed embryo transfer on birthweight, macrosomia, and low birthweight rates in US singleton infants. Am. J. Obstet. Gynecol. 218,. https://doi.org/10.1016/j.ajog.2017.12.223 433.e1-433.e10.

Mackens, S., Santos-Ribeiro, S., van de Vijver, A., Racca, A., Van Landuyt, L., Tournaye, H., Blockeel, C., 2017. Frozen embryo transfer: a review on the optimal endometrial preparation and timing. Hum. Reprod. 32, 2234–2242. https://doi.org/10.1093/humrep/dex285.

Mumusoglu, S., Polat, M., Ozbek, I.Y., Bozdag, G., Papanikolaou, E.G., Esteves, S.C., Humaidan, P., Yarali, H., 2021. Preparation of the Endometrium for Frozen Embryo Transfer: A Systematic Review. Front. Endocrinol. 12, 688237. https://doi.org/10.3389/fendo.2021.688237.

Pereira, M.M., Mainigi, M., Strauss, J.F., 2021. Secretory products of the corpus luteum and preeclampsia. Hum. Reprod. Update 27, 651–672. https://doi.org/10.1093/humupd/dmab003.

Saito, K., Kuwahara, A., Ishikawa, T., Morisaki, N., Miyado, M., Miyado, K., Fukami, M., Miyasaka, N., Ishihara, O., Irahara, M., Saito, H., 2019. Endometrial preparation methods for frozen-thawed embryo transfer are associated with altered risks of hypertensive disorders of pregnancy, placenta accreta, and gestational diabetes mellitus. Hum. Reprod. 34, 1567–1575. https://doi.org/10.1093/humrep/dez079.

von Versen-Höynck, F., Griesinger, G., 2022. Should any use of artificial cycle regimen for frozen-thawed embryo transfer in women capable of ovulation be abandoned: yes, but what's next for FET cycle practice and research? Hum. Reprod. 37, 1697–1703. https://doi.org/10.1093/humrep/deac125.

von Versen-Höynck, F., Schaub, A.M., Chi, Y.-Y., Chiu, K.-H., Liu, J., Lingis, M., Stan Williams, R., Rhoton-Vlasak, A., Nichols, W.W., Fleischmann, R.R., Zhang, W., Winn, V.D., Segal, M.S., Conrad, K.P., Baker, V.L., 2019. Increased Preeclampsia Risk and Reduced Aortic Compliance With In Vitro Fertilization Cycles in the Absence of a Corpus Luteum. Hypertension 73, 640–649. https://doi.org/10.1161/HYPERTENSIONAHA.118.12043.

Weiss, A., Baram, S., Geslevich, Y., Goldman, S., Nothman, S., Beck-Fruchter, R., 2021. Should the modified natural cycle protocol for frozen embryo transfer be modified? A prospective case series proof of concept study. Eur. J. Obstet. Gynecol. Reprod. Biol. 258, 179–183. https://doi.org/10.1016/j.ejogrb.2021.01.004.

Wolfe, E.L., Vaughan, D., Craig, W., Amaral, B., Penzias, A., Sakkas, D., Toth, T.L., 2023. Modified natural and optimized programmed frozen embryo transfers have equivalent live birth rates: an analysis of 6,682 cycles. Fertil. Steril. 120, 80–88. https://doi.org/10.1016/j.fertnstert.2023.02.020.

Zaat, T., De Bruin, J.-P., Goddijn, M., Van Baal, M., Benneheij, S., Brandes, M., Broekmans, F., Cantineau, A., Cohlen, B., Van Disseldorp, J., Gielen, S., Groenewoud, E., Van Heusden, A., Kaaijk, E., Koks, C., De Koning, C., Klijn, N., Van Der Linden, P., Manger, P., Moolenaar, L., Van Oppenraaij, R., Pieterse, Q., Smeenk, J., Visser, J., Van Wely, M., Mol, F., 2023. Home-based monitoring of ovulation to time frozen embryo transfers in the Netherlands (Antarctica-2): an open-label, nationwide, randomised, non-inferiority trial. The Lancet 402, 1347–1355. https://doi.org/10.1016/S0140-6736(23)01312-0.

Received 29 September 2023; received in revised form 18 November 2023; accepted 23 November 2023.

RBMO

VIEWPOINT

Avoiding weekend frozen embryo transfer in modified natural cycles: is it possible?

Barbara Lawrenz[1,2], Christophe Blockeel[3,4,*]

ABSTRACT

In this era of the freeze-all strategy, the prevalence of frozen embryo transfer (FET) cycles is increasing rapidly. Although still quite often used, the hormone replacement therapy cycle to prepare a FET should now belong to the past, unless strictly necessary. This raises questions about possible flexible protocols for the preparation of an FET cycle in a (modified) natural cycle. In this viewpoint, an overview of the different options is discussed, stressing the importance of the corpus luteum.

The shift from fresh towards frozen embryo transfer (FET) cycles seems to be irreversible, with FET cycles currently being performed on a daily basis in many IVF centres across the globe. Given the magnitude of the increase in FET cycles, consideration needs to be given not only to the question of the impact of the endometrial preparation approach on the implantation/pregnancy rate, the course of pregnancy and the health of the mother and offspring, but also to the ease of planning and performing the FET with all the steps required.

Endometrial preparation for FET cycles can basically be performed as a medicated cycle (hormonal replacement therapy [HRT]) or as a 'pure' natural cycle (pNC; without any intervention) or 'modified' natural cycle (mNC; with or without hormonal stimulation and with the administration of human chorionic gonadotrophin [HCG] for ovulation induction). In an HRT cycle, the administration of oestrogen initiates endometrial proliferation. Serum oestrogen concentrations exhibit negative feedback on the hypothalamic–pituitary–ovarian axis, suppressing follicle growth, the development of a dominant follicle, ovulation and consequently also the formation of a corpus luteum. Evidence is mounting that pregnancy complications (bleeding, miscarriage, pregnancy-induced hypertension and pre-eclampsia) (*Roelens and Blockeel, 2022*) are more frequent in pregnancies conceived through HRT (HRT–FET). This is likely to be due to the absence of a corpus luteum, which results in a lack of production of relaxin and other vasoactive substances that are important for the mother's adaptation to the pregnancy (*Roelens and Blockeel, 2022*).

Contemporary evidence consequently favours the natural-cycle approach, which allows follicle growth and corpus luteum formation, thereby presenting the optimal strategy for mitigating the hazards associated with the HRT approach. The challenge in planning FET in a pNC arises from accurately identifying ovulation with the subsequent progesterone surge, initiating the secretory transformation of the endometrium. The endometrium is thought to be receptive about 120 h after the progesterone rise begins, and the window of implantation allows the implantation of embryos for a period of about 2 days (*Doyle et al., 2022*).

Since secretory transformation of the endometrium is not directly assessable, urinary LH, serum LH and serum progesterone serve as a proxy for the endometrial changes, initiated by ovulation and the progesterone rise. These hormones are commonly used to time embryo transfers correctly. Urinary LH kits for ovulation detection present a patient-friendly, convenient and low-cost approach, and a recently published randomized controlled trial (*Zaat et al., 2021*) demonstrated the 'non-inferiority' of home-based urinary LH testing versus hospital-controlled HCG triggering in terms of the ongoing pregnancy rate. However, as pointed out by Mackens and colleagues (*Mackens and Blockeel, 2023*), these results require careful consideration and additional elaboration, and urinary LH detection may not be the most accurate method for ovulation detection.

The restriction of the sole use of serum LH samples in determining the onset of the subsequent progesterone rise is the existing variation in the time intervals between the LH peak and the rise in progesterone (*Coughlan et al., 2023*). Ideally, hormonal measurements for the correct identification of ovulation and the progesterone rise in a pNC should include measurement of LH, oestrogen and progesterone (*Lawrenz et al., 2023*). However, repeated measurements of

[1] ART Fertility Clinics, Abu Dhabi, UAE
[2] Reproductive Unit, UZ Ghent, Belgium
[3] Brussels IVF, Universitair Ziekenhuis Brussel, Brussels, Belgium
[4] Department of Obstetrics and Gynaecology, University of Zagreb School of Medicine, Zagreb, Croatia

*Corresponding author. E-mail address: Christophe.Blockeel@uzbrussel.be (C. Blockeel). https://doi.org/10.1016/j.rbmo.2023.103775 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS
Frozen embryo transfer
Modified natural cycle
Natural cycle

these three hormones induce a higher physical and financial burden on patients and clinics as this requires frequent clinic visits. Furthermore, the possible 'untimely' occurrence of ovulation might require an embryo transfer procedure to be performed on a weekend (e.g. a Sunday) or a holiday, days on which the full workforce of the IVF clinic might not be regularly present in the clinic.

The ideal scenario would eventually be presented by the existence of a corpus luteum along with the possibility for improved scheduling of the embryo transfer procedure, and therefore the workload. The mNC approach only partially combines these two factors as the trigger is commonly administered at a follicle size around 17 mm, which again limits the freedom of ET procedure planning.

To 'free' oneself from this obstacle and facilitate planning flexibility, the study of Alonso-Mayo and colleagues (*Alonso-Mayo et al., 2023*) evaluated the impact of triggering at different follicle sizes on the ongoing pregnancy rate in mNC cycles. Follicle sizes at the time of triggering were stratified into three groups (13.0–15.9 mm, 16.0–18.9 mm and ≥19.0 mm). Whereas the initial analysis showed a significant difference between the groups, favouring trigger at smaller follicle sizes, the significance disappeared when the analysis was adjusted for confounders of preimplantation genetic testing for aneuploidies and oocyte donation.

As the authors did not see a difference in the ongoing pregnancy rates, they concluded that from a follicle size of 13 mm onwards (taking into account an assumed follicular growth rate of 1–1.5 mm/day), triggering between 13 and 22 mm is possible. This approach would allow a flexible range of 5–7 days for FET planning in patients undergoing an mNC protocol, thereby facilitating embryo transfer planning for both patients and IVF clinics. Moreover, the protective effects of the corpus luteum beyond progesterone production are present. Hence, the authors point to the fact that a progesterone rise of over 1.5 ng/ml must be excluded on the trigger day.

An 'HCG-free natural cycle approach' is the so-called natural proliferative-phase protocol, which would also allow a certain freedom of scheduling. Exogeneous progesterone is started once a follicle of at least 14 mm is present, and hence the spontaneous LH surge has not yet started. Initial data point to similar clinical outcomes and the added progesterone might even trigger ovulation (*Santos-Ribeiro et al., 2023*), thereby maintaining the benefit of a corpus luteum throughout the course of pregnancy.

However, for both approaches, which allow more freedom in the scheduling of FET, further research is warranted to identify whether the corpus luteum, which results from smaller follicle sizes, produces enough relaxin and vasoactive substances to achieve the same protective effects as a follicle, which ovulates naturally in a pNC.

REFERENCES

Alonso-Mayo, C., Kohls, G., Santos-Ribeiro, S., Soares, S.R., Garcia-Velasco, J.A., 2023. Modified natural cycle allows a window of seven days for frozen embryo transfer planning. Reprod Biomed Online 103774. https://doi.org/10.1016/j.rbmo.2023.103774.

Coughlan, C., Ata, B., Gallego, R.D., Lawrenz, B., Melado, L., Samir, S., Fatemi, H., 2023. Interindividual variation of progesterone elevation post LH rise: implications for natural cycle frozen embryo transfers in the individualized medicine era. Reprod. Biol. Endocrinol. 21, 47. https://doi.org/10.1186/s12958-023-01096-4.

Doyle, N., Jahandideh, S., Hill, M.J., Widra, E.A., Levy, M., Devine, K., 2022. Effect of Timing by Endometrial Receptivity Testing vs Standard Timing of Frozen Embryo Transfer on Live Birth in Patients Undergoing In Vitro Fertilization: A Randomized Clinical Trial. JAMA 328, 2117–2125. https://doi.org/10.1001/jama.2022.20438.

Lawrenz, B., Melado, L., Fatemi, H.M., 2023. Frozen embryo transfers in a natural cycle: how to do it right. Current Opinion in Obstetrics & Gynecology Publish Ahead of Print. https://doi.org/10.1097/GCO.0000000000000862.

Mackens, S., Blockeel, C., 2023. Home-based monitoring prior to frozen embryo transfer: the new gold standard? The Lancet 402, 1304–1306. https://doi.org/10.1016/S0140-6736(23)01798-1.

Roelens, C., Blockeel, C., 2022. Impact of different endometrial preparation protocols before frozen embryo transfer on pregnancy outcomes: a review. Fertility and Sterility 118, 820–827. https://doi.org/10.1016/j.fertnstert.2022.09.003.

Santos-Ribeiro, S., Godinho, C.M., Reis-Soares, S., 2023. Nature (almost) always prevails — challenging the status quo of artificial cycle frozen embryo transfers. Reproductive BioMedicine Online 47, 103352. https://doi.org/10.1016/j.rbmo.2023.103352.

Zaat, T.R., de Bruin, J.P., Goddijn, M., van Baal, M., Benneheij, E.B., Brandes, E.M., Broekmans, F., Cantineau, A.E.P., Cohlen, B., van Disseldorp, J., Gielen, S.C.J.P., Groenewoud, E.R., van Heusden, A., Kaaijk, E.M., Koks, C., de Koning, C.H., Klijn, N.F., Lambalk, C.B., van der Linden, P.J.Q., Manger, P., van Oppenraaij, R.H.F., Pieterse, Q., Smeenk, J., Visser, J., van Wely, M., Mol, F., 2021. Is home-based monitoring of ovulation to time frozen embryo transfer a cost-effective alternative for hospital-based monitoring of ovulation? Study protocol of the multicentre, non-inferiority Antarctica-2 randomised controlled trial. Human Reproduction Open 2021, hoab035. https://doi.org/10.1093/hropen/hoab035.

Received 29 November 2023; accepted 1 December 2023.

RBMO

ARTICLE

Impact of aneuploidy on reproductive success in young infertile women: prospective analysis

BIOGRAPHY

Mandy Katz-Jaffe is a reproductive geneticist who completed her Masters and PhD degrees at Monash University, Australia. After receiving her doctorate, she joined CCRM Fertility in the USA, where she is the Scientific and Genetics Director. Dr Katz-Jaffe has extensive clinical expertise in PGT and oversees an active reproductive biology research group.

Mandy Katz-Jaffe*, Carly Gassen, Rachel Makloski, Laura Reed, William B. Schoolcraft

KEY MESSAGE

Young infertile women have significantly improved live birth potential through embryo selection using preimplantation genetic testing for aneuploidies (PGT-A) with their first attempt at conception. From the point of the initial physician consultation, incorporating PGT-A as an embryo selection method resulted in the fastest path to the healthiest singleton baby.

ABSTRACT

Research question: What is the clinical outcome of the first attempt at conception between two embryo selection methods, blastocyst morphology and preimplantation genetic testing for aneuploidies (PGT-A), chosen at the initial physician IVF consultation?

Design: In this prospective analysis, a clinical decision regarding embryo selection, blastocyst morphology (group A) or PGT-A (group B) was made during initial physician IVF consultation. Female infertility patients were matched based on maternal age (mean 32.6 ± 3.6 years; range 25–43 years) and a similar time frame of oocyte retrieval. The primary outcome was live birth rate from the initial consultation to the first conception attempt for all female patients and for a subset analysis of patients aged <35 years.

Results: The inclusion of PGT-A (group B) for embryo selection during the initial physician IVF consultation resulted in 23 additional women out of the total 100 achieving a healthy live birth following the first conception attempt in this maternally age-matched infertile population (group B = 72.0% versus group A = 49.0%; $P = 0.0014$). This same benefit was observed for age-matched, younger infertility patients (<35 years), with live birth rates from the initial consultation being significantly higher when the upfront clinical decision included PGT-A for embryo selection (group B = 76.7% versus group A = 53.4%; $P = 0.0052$). Interestingly, 17 women from group B would have received an aneuploid embryo transfer if selection had been determined by blastocyst morphology alone, as their best-grade embryo was aneuploid.

Conclusions: This prospective analysis from the initial physician IVF consultation revealed that euploid embryo selection significantly improved live birth potential with the first conception attempt, even for younger women with infertility.

INTRODUCTION

Fetal aneuploidy is a common, natural occurrence that is associated with higher rates of implantation failure, miscarriage and birth defects (*Demko et al., 2016*). An estimated 50–60% of first-trimester pregnancy losses occur due to chromosomal aneuploidy, which is significantly associated with maternal age (*Hodes-Wertz et al., 2012*; *Werner et al., 2012*). This is also observed in preimplantation embryos, with infertile women younger than 30 years exhibiting less than 30% aneuploidy, increasing to a rate of over 85% for women in their mid-40s (*Franasiak et al., 2014*).

CCRM Fertility, Lone Tree, Colorado, USA.

The abstract of this research article was presented at the American Society for Reproductive Medicine (ASRM) 2023 Scientific Congress and Expo, which took place on 14–18 October 2023 in New Orleans, Louisiana, USA.

*Corresponding author. E-mail address: MandyK@colocrm.com (M. Katz-Jaffe). https://doi.org/10.1016/j.rbmo.2024.103858

Declaration: The authors report no financial or commercial conflicts of interest.

KEYWORDS

Aneuploidy
Frozen embryo transfer
Infertility
IVF
Preimplantation genetic testing

Preimplantation embryonic morphology has only a weak association with aneuploidy, necessitating a molecular analysis of biopsied embryonic cells to determine the ploidy, referred to as preimplantation genetic testing for aneuploidies (PGT-A) (*Scheffer et al., 2023*).

Randomized clinical trials (RCT) and a recent meta-analysis have demonstrated improved IVF outcomes with the use of PGT-A for embryo selection when reporting whole-chromosome aneuploidy, including an increased implantation rate, ongoing pregnancy rate (OPR) and live birth rate (*Dahdouh et al., 2015*; *Kasaven et al., 2023*; *Lee et al., 2019*; *Neal et al., 2018*; *Pantou et al., 2022*; *Rubio et al., 2017*). PGT-A has also been shown to reduce treatment cost and shorten the time to pregnancy (*Neal et al., 2018*; *Rubio et al., 2017*). Although these results are promising, there is debate in the literature regarding the benefit of PGT-A as an embryo selection method for women less than 38 years of age. Indeed, a few RCT have indicated that PGT-A might be interpreted as an impairment or offer no improvement for younger infertile women (*Munne et al., 2019*; *Ozgur et al., 2019*; *Yan et al., 2021*). Different strategies for randomization, experimental design and the interpretation and decision making related to presumptive mosaic embryos could all represent contributing variables.

To most accurately assess the effectiveness of PGT-A as an embryo selection strategy, an intention to treat (ITT) analysis was performed. The objective of this prospective study was to compare clinical IVF outcomes following the first attempt at conception between two embryo selection methods, blastocyst morphology and PGT-A, chosen at the initial physician IVF consultation.

MATERIALS AND METHODS

This prospective, ITT analysis was performed from 2016 to 2021, evaluating de-identified data of infertility patients at a single private infertility clinic following their initial physician IVF consultation. During this consultation, a decision regarding embryo selection was agreed upon between the reproductive endocrinologist (REI) and the patient, choosing blastocyst morphology or PGT-A. Once an ovarian stimulation calendar had been created, female patients were matched based on maternal age and a similar time frame of oocyte retrieval.

Group A consisted of 100 patients who were physician directed to an IVF cycle with embryo selection based on blastocyst morphology. Group B consisted of 100 patients who were physician directed to an IVF cycle with embryo selection based on PGT-A. Each patient received thorough counselling and procedural information, as well as a genetic counselling consultation for the PGT-A group, alongside standard informed consent for IVF and frozen embryo transfer (FET). This study was reviewed by the Western Institutional Review Board on 20 August 2018 and was determined to be exempt because it involved a secondary data analysis of de-identified data and did not include experimental interventions (WIRB Work Order #1-1103486-1).

All patients underwent routine ovarian stimulation individualized to each patient by the REI based on maternal age, ovarian reserve and prior responses to ovarian stimulation, if applicable. Ovarian reserve was measured by anti-Müllerian hormone (AMH), day 3 FSH and day 3 oestradiol concentrations, and antral follicle count (AFC). Routine transvaginal ultrasound-guided oocyte retrieval was performed with mature metaphase II (MII) oocytes fertilized through intracytoplasmic sperm injection, in line with the standard clinic protocol. Embryos were cultured in sequential media (SAGE; CooperSurgical Fertility Solutions, USA) and grown until the blastocyst stage of development. All blastocysts were graded using the Gardner–Schoolcraft scoring system, with modification to account for early hatching of the blastocysts (*Schiewe et al., 2015*; *Schoolcraft et al., 2011*). For all patients, blastocysts were vitrified according to routine laboratory protocols and an established Cryotop method Kitazato, Japan (*Gardner and Schoolcraft, 1999*).

For patients in group B, trophectoderm biopsy and PGT-A were performed. On day 3 of embryonic development, an opening (3–5 μm) in the zona pellucida was created using a series of five laser pulses at 100% power for 200 μs each (Hamilton Thorne Research, USA). Once an identifiable inner cell mass was observed in blastocyst-stage embryos (5–6 days after oocyte retrieval), a trophectoderm biopsy was performed by laser dissection as previously described, and the biopsied blastocysts were vitrified using the previously established Cryotop method (*Gardner and Schoolcraft, 1999*). PGT-A was performed on the biopsied trophectoderm cells using NGS VeriSeq (Vitrolife, Sweden) with ploidy prediction using BlueFuse software (Illumina, USA) (*Munne et al., 2019*). In addition to NGS QC, a euploid classification was defined at 0–20%, low presumptive mosaic at 20–50%, high presumptive mosaic at 50–80%, and aneuploid classification at 80–100%.

The standard clinical protocol for endometrial preparation using hormone replacement was employed (*Surrey et al., 2017*). Vitrified blastocysts were warmed on the day of transfer and transferred under ultrasound guidance (*Surrey et al., 2018*). Nine days after the embryo transfer, each patient had a serum pregnancy test. In the event of a positive pregnancy test, serum human chorionic gonadotrophin (HCG) concentrations were serially tested to ensure that HCG concentrations were adequately rising. Confirmation of clinical pregnancy with fetal heart tones (FHT) via ultrasound examination was performed within 4 weeks of the increasing serum HCG concentrations.

The primary outcome of this prospective study was to compare the live birth following the first attempt at conception, based on the embryo selection method (blastocyst morphology or PGT-A) that was chosen at the initial physician IVF consultation. Other study outcomes included implantation with FHT, miscarriage and live birth rates from FET. The rate of implantation with FHT was defined as the rate of transferred embryos that demonstrated clinical pregnancy confirmation from FHT on ultrasound examination at 6.5 weeks' gestation. The clinical miscarriage rate was defined as the proportion of transferred embryos that resulted in a pregnancy loss following the confirmation of viable implantation with FHT. The live birth rate was calculated from two starting points: the live birth rate per embryo transfer and live birth rate per IVF cycle start. Ovarian reserve parameters (day 3 oestradiol, day 3 FSH, AMH and AFC) and IVF cycle results (number of oocytes retrieved, number of mature MII oocytes fertilized, and number of day 5 and day 6 blastocysts) were measured for both groups.

The outcome data, ovarian reserve parameters and IVF cycle results were compared between group A and group B

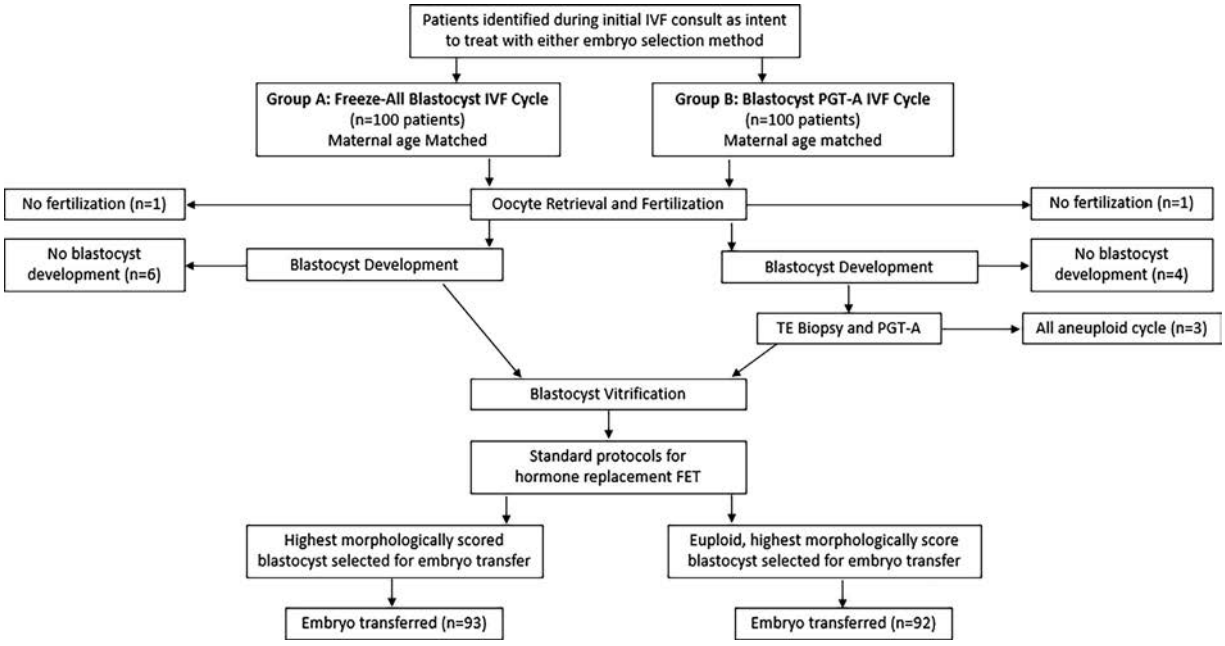


FIGURE 1 Assignment and treatment of female patients with infertility. FET, frozen embryo transfer; PGT-A, preimplantation genetic testing for aneuploidies; TE, trophectoderm.

using a two-sided Fisher's exact test or unpaired two-sided t-test, as appropriate. The data were base 10 log-transformed prior to the two-sided t-test. A *P*-value <0.05 was considered statistically significant. All data analysis was performed using Prism 8 software (GraphPad, USA).

RESULTS

The female patients were maternally age matched (mean age 32.6 ± standard deviation 3.6 years; range 25–43 years) into two groups (group A, blastocyst morphology; group B, PGT-A; $n = 100$ per group) based on the embryo selection method chosen during the initial physician IVF consultation. The number of patient cycles resulting in no oocyte fertilization or no blastocyst development was not significantly different between the two groups (Supplementary Table 1). The results for each group included one patient cycle with no oocyte fertilization. Six patient cycles in group A and four patient cycles in group B resulted in early embryo arrest with no blastocyst development. In group B, three patient cycles were identified with only aneuploid embryos, based on the PGT-A results.

A total of 93 patient cycles in group A and 92 patient cycles in group B went forward with an FET (FIGURE 1). The ovarian reserve, measured by day 3 oestradiol, day 3 FSH, AMH and AFC, did not differ significantly between groups A and B (TABLE 1A, Supplementary Table 1). There were no statistically significant differences in the IVF cycle parameters (number of oocytes retrieved, number of mature MII oocytes fertilized, and number and grade of day 5 and day 6 blastocysts) between the two groups (TABLE 2A, Supplementary Table 1). Slower developing embryos were cultured to day 7 of embryonic development. Only one patient cycle in group A and 12 patient cycles in group B resulted in slower day 7 blastocysts. For group B cycles with PGT-A for embryo selection, the results revealed 64.1% (396/618) euploid, 33.5% (207/618) aneuploid (chromosome errors illustrated in Supplementary Figure 1) and 2.4% (15/618) mosaic blastocysts. In all group B cycles with a presumptive mosaic embryo there was a euploid blastocyst with higher morphological grade available for transfer.

Clinical outcomes were significantly improved when the decision for embryo selection during the initial physician IVF consultation was PGT-A (group B) over blastocyst morphology (group A; TABLE 3). There was no significant difference in the blastocyst grade or number of embryos transferred between the two maternally age-matched groups (Supplementary Table 1). The implantation rate with FHT was significantly higher in group B than group A (76.8% versus 50.8% respectively; $P < 0.0001$; TABLE 3), with 11 sets of twins observed in group B and 7 sets of twins in group A. A trend towards a lower miscarriage rate was observed in group B (4.0% versus 10.9% respectively; TABLE 3), although this was not statistically significant. Both the live birth rate per embryo transfer (group B = 78.3% versus group A = 52.7%; $P = 0.0003$) and live birth rate per initial physician IVF consultation (group B = 72.0% versus group A = 49.0%; $P = 0.0014$) were significantly higher with group B and the embryo selection method of PGT-A (TABLE 3). The upfront clinical decision during the initial physician IVF consultation to include PGT-A (group B) for embryo selection in this maternally age-matched infertile population resulted in 23 additional women reaching a live birth.

Data analyses were also performed for only younger female patients <35 years of age; there were 73 maternally age-matched patients in each group (mean age 30.8 ± 2.1 years; range 25–34 years) with a similar time frame of oocyte retrieval. In group A, two patient cycles resulted in early embryo arrest with no blastocyst development. Comparably, in group B, one patient cycle resulted in no oocyte fertilization and one patient cycle resulted in early embryo arrest with no blastocyst development. All the cycles in younger female patients in group B resulted in at least one euploid blastocyst with 65.9% (331/502) euploid, 31.5% (158/502)

TABLE 1 OVARIAN RESERVE DATA

(A) All infertility patients (mean age 32.6 ± 3.6 years)

Parameter	Group A: freeze-all IVF cycle (*n* = 100)	Group B: PGT-A IVF cycle (*n* = 100)	*P*-value
Day 3 oestradiol (pg/ml)	41.3 ± 16.6	41.6 ± 18.3	0.93
Day 3 FSH (mIU/ml)	7.1 ± 2.1	7.5 ± 2.0	0.19
AMH (ng/ml)	5.0 ± 5.8	3.8 ± 2.8	0.48
AFC (*n*)	26.7 ± 17.3	21.8 ± 12.2	0.13

(B) Infertility patients under age 35 years (mean age 30.8 ± 2.1 years)

Parameter	Group A: freeze-all IVF cycle (*n* = 73)	Group B: PGT-A IVF cycle (*n* = 73)	*P*-value
Day 3 oestradiol (pg/ml)	43.7 ± 16.9	40.8 ± 18.6	0.09
Day 3 FSH (mIU/ml)	6.7 ± 1.7	7.3 ± 2.0	0.11
AMH (ng/ml)	5.6 ± 6.2	4.1 ± 3.0	0.13
AFC (*n*)	29.7 ± 17.3	23.5 ± 12.8	0.05

Data are shown as mean value ± standard deviation.

AFC, antral follicle count; AMH, anti-Müllerian hormone; PGT-A, preimplantation genetic testing for aneuploidies.

TABLE 2 IVF CYCLE RESULTS

(A) All infertility patients (mean age 32.5 ± 3.6 years)

Parameter	Group A: freeze-all IVF cycle (*n* = 100)	Group B: PGT-A IVF cycle (*n* = 100)	*P*-value
Number of oocytes retrieved	23.2 ± 12.5	20.7 ± 9.2	0.11
Number of MII oocytes fertilized	12.5 ± 7.6	12.3 ± 5.9	0.41
Number of day 5 blastocysts	3.9 ± 4.6	3.0 ± 2.7	0.59
Number of day 6 blastocysts	2.5 ± 2.1	3.3 ± 2.5	0.06

(B) Infertility patients under age 35 years (mean age 30.8 ± 2.1 years)

Parameter	Group A: freeze-all IVF cycle (*n* = 73)	Group B: PGT-A IVF cycle (*n* = 73)	*P*-value
Number of oocytes retrieved	25.1 ± 12.1	21.4 ± 9.2	0.15
Number of MII oocytes fertilized	13.5 ± 7.7	12.8 ± 6.1	0.82
Number of day 5 blastocysts	4.4 ± 4.9	3.5 ± 2.7	0.41
Number of day 6 blastocysts	2.7 ± 2.2	3.5 ± 2.5	0.11

Data are shown as mean value ± standard deviation.

MII, metaphase II; PGT-A, preimplantation genetic testing for aneuploidies.

TABLE 3 CLINICAL OUTCOMES FOR ALL INFERTILITY PATIENTS (MEAN AGE 32.5 ± 3.6 YEARS)

Parameter	Group A (*n* = 71 FET) (%, *n*/*N*)	Group A (*n* = 71 FET) (%, *n*/*N*)	*P*-value with RR and 95% CI
Implantation rate (FHT)	50.8 (61/120)	76.8 (86/112)	<0.0001 RR 0.5978; 95% CI 0.4711–0.787
Miscarriage rate	10.9 (6/55)	4.0 (3/75)	0.1668
Live birth rate per FET	52.7 (49/93)	78.3 (72/92)	0.0003 RR 0.5890; 95% CI 0.4488–0.7731
Live birth rate per initial physician IVF consultation	49.0 (49/100)	72.0 (72/100)	0.0014 RR 0.6273; 95% CI 0.4784–0.8225

Group A, embryo selection based on morphology; group B, embryo selection based on PGT-A.

FET, frozen embryo transfer; FHT, fetal heart tones; PGT-A, preimplantation genetic testing for aneuploidies; RR, Relative risk.

TABLE 4 CLINICAL OUTCOMES FOR INFERTILITY PATIENTS UNDER AGE 35 YEARS (MEAN AGE 30.8 ± 2.1 YEARS)

Parameter	Group A (*n* = 71 FET) (%, *n*/*N*)	Group B (*n* = 71 FET) (%, *n*/*N*)	*P*-value with RR and 95% CI
Implantation rate (FHT)	52.2 (48/92)	77.9 (67/86)	0.0005 RR 0.5976; 95% CI 0.4561–0.7830
Miscarriage rate	9.3 (4/43)	3.5 (2/58)	0.3972
Live birth rate per FET	54.9 (39/71)	78.9 (56/71)	0.0041 RR 0.6030; 95% CI 0.4420–0.8225
Live birth rate per initial physician IVF consultation	53.4 (39/73)	76.7 (56/73)	0.0052 RR 0.6158; 95% CI 0.4519–0.8391

Group A, embryo selection based on morphology; group B, embryo selection based on PGT-A.

FET, frozen embryo transfer; FHT, fetal heart tones; PGT-A, preimplantation genetic testing for aneuploidies; RR, Relative risk.

aneuploid and 2.6% (13/502) mosaic embryos. A total of 71 patient cycles in both groups went forward with an FET. Ovarian reserve, measured by day 3 oestradiol, day 3 FSH, AMH and AFC, did not differ significantly between groups A and B (TABLE 1B). There were no statistically significant differences in IVF cycle parameters (number of oocytes retrieved, number of mature MII oocytes fertilized or number of day 5 and day 6 blastocysts) between the two groups (TABLE 2B).

Clinical outcomes for younger patients with infertility were significantly improved when the decision for embryo selection during the initial physician IVF consultation was PGT-A (group B) over blastocyst morphology (group A; TABLE 4). There was no significant difference in the number of embryos transferred between the two maternally age-matched groups (Supplementary Table 1). The implantation rate with FHT was significantly higher in group B than group A (77.9% versus 52.2% respectively; $P = 0.0005$; TABLE 4), with nine sets of twins observed in group B and five sets of twins in group A. A trend towards a lower miscarriage rate was observed in group B (3.5% versus 9.3% respectively; TABLE 4), although this was not statistically significant. Both the live birth rate per embryo transfer (group B = 78.9% versus group A = 54.9%; $P = 0.0041$) and live birth rate per initial physician IVF consultation (group B = 76.7% versus group A = 53.4%; $P = 0.0052$) were significantly higher with group B and the embryo selection method of PGT-A (TABLE 4). The upfront clinical decision during the initial physician IVF consultation to include PGT-A (group B) for embryo selection in this maternally age-matched young (<35 years) infertile population resulted in 17 additional women reaching a live birth.

DISCUSSION

For individuals with infertility, the goal of IVF is to achieve a healthy live delivery in the shortest treatment time. Selecting the most competent embryo of the patient's available cohort is critical to maximize IVF success and live birth potential. This prospective ITT analysis compared the first live birth outcome between two embryo selection methods, embryo morphology and PGT-A. This decision was REI directed alongside patient consent during the initial physician IVF consultation.

The results of this prospective analysis demonstrated that when embryo selection with PGT-A was chosen before treatment initiation during a new patient REI consultation, infertility patients including women <35 years saw clinical outcomes significantly improved, with higher rates of implantation and healthy live birth. The decision to use embryo selection with PGT-A ultimately resulted in an impressive 23 additional infertile women delivering healthy infants from the initial 100 women who began treatment in the PGT-A embryo selection group, upon their first attempt at conception. Interestingly, 17 of these women would have received an aneuploid embryo transfer if selection had been determined by blastocyst morphology in place of PGT-A as their best-grade embryo was indeed chromosomally abnormal.

Both the live birth rate per initial physician IVF consultation and live birth rate per embryo transfer were significantly increased following embryo selection with PGT-A in this cohort of infertility patients and the subset of younger infertility patients (<35 years). In contrast, a few RCT in the literature have indicated that PGT-A might be interpreted as an impairment for younger infertile women (*Munne et al., 2019*; *Ozgur et al., 2019*; *Yan et al., 2021*). This was specifically noted in the Single Embryo TrAnsfeR of Euploid Embryo (STAR) trial by Munné and colleagues, which found no significant benefit in 20-week OPR per FET or IVF cycle start in women aged 25–40 years, with younger infertile women (<35 years) in the study even showing a decrease in OPR per FET with PGT-A (*Munne et al., 2019*). Additionally, the Society for Assisted Reproductive Technology Clinic Outcome Reporting System (SART CORS) database retrospective analysis has shown no improvement in cumulative live birth rate per cycle start for women aged less than 35 years (*Kucherov et al, 2023*; *Mejia et al, 2022*).

This variability in clinical outcomes may be explained by differences in the study populations, such as maternal age and infertility indication, as well as differences in individual fertility clinic protocols and laboratory techniques (IVF and genetic). Indeed, Bardos and co-workers provide an insight into the role of the PGT-A laboratory and the variable outcomes across PGT-A laboratories, highlighting differences in techniques and bioinformatics, with the use of next-generation sequencing technologies (*Bardos et al., 2023*). Non-selection studies evaluating pregnancy outcomes after a blinded transfer of euploid versus aneuploid embryos have also demonstrated a near-perfect negative predictive value for an aneuploid result to predict live birth (*Tiegs et al., 2021*). Using

PGT-A to identify euploid embryos for transfer, while simultaneously selecting against aneuploid embryos, minimizes the burden of IVF by reducing futile transfers, miscarriages and conceptions with fetal aneuploidy (*Hynes and Forman, 2023*).

The results of this prospective, initial physician IVF consultation, ITT analysis indicate that embryo selection with PGT-A maximizes the potential of reproductive success and live birth from the first attempt at conception, even for younger women with infertility for whom IVF is indicated. These findings are critical for any individual with infertility pursuing IVF, as they suggest that from the initial physician consultation the decision to include PGT-A for embryo selection is a crucial component of the path to reach a healthy delivery in the shortest treatment time. A major strength of this study was the upfront clinical decision ITT analysis, which better portrays pragmatic IVF outcomes in a fertility clinic setting by accounting for every patient who initiated treatment. PGT-A chosen for embryo selection ultimately improved clinical outcomes at the first attempt at conception, including for the younger infertile patient population in the absence of advanced maternal ageing and the associated significantly higher oocyte aneuploidy rates.

In conclusion, this study showed that the upfront embryo selection strategy chosen at the initial physician IVF consultation significantly impacts reproductive success and the probability of a live birth at the first attempt at conception. Even within a population of younger, infertile women for whom IVF is indicated, live birth potential was maximized with embryo selection by PGT-A, over blastocyst morphology alone. The incorporation of PGT-A as an embryo selection strategy for younger patients with infertility, with the ability to identify and transfer a euploid embryo, resulted in the fastest path to the healthiest singleton baby.

DATA AVAILABILITY

The data that has been used is confidential.

ACKNOWLEDGEMENTS

The authors acknowledge and thank the PGT laboratory team, IVF lab laboratory team, genetic counsellors, physicians, nurses and staff at CCRM Colorado. The authors also thank the IVF patients at CCRM Colorado, without whom this research could not have been conducted.

ETHICS DECLARATION

This study was reviewed by the Western Institutional Review Board and was determined to be exempt because it involved a secondary data analysis of de-identified data and did not include experimental interventions.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.rbmo.2024.103858.

REFERENCES

Bardos, J., Kwal, J., Caswell, W., Jahandideh, S., Stratton, M., Tucker, M., DeCherney, A., Devine, K., Hill, M., 2023. Reproductive genetics laboratory may impact euploid blastocyst and live birth rates: a comparison of 4 national laboratories' PGT-A results from vitrified donor oocytes. Fertil. Steril. 119 (1), 29–35. https://doi.org/10.1016/j.fertnstert.2022.10.010.

Dahdouh, E.M., Balayla, J., García-Velasco, J.A., 2015. Comprehensive chromosome screening improves embryo selection: a meta-analysis. Fertil. Steril. 104 (6), 1503–1512. https://doi.org/10.1016/j.fertnstert.2015.08.038.

Demko, Z.P., Simon, A.L., McCoy, R.C., Petrov, D.A., Rabinowitz, M., 2016. Effects of maternal age on euploidy rates in a large cohort of embryos analyzed with 24-chromosome single-nucleotide polymorphism–based preimplantation genetic screening. Fertil. Steril. 105 (5), 1307–1313. https://doi.org/10.1016/j.fertnstert.2016.01.025.

Franasiak, J.M., Forman, E.J., Hong, K.H., Werner, M.D., Upham, K.M., Treff, N.R., Scott, Jr., R.T., 2014. The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertil. Steril. 101 (3), 656–663.e1. https://doi.org/10.1016/j.fertnstert.2013.11.004.

Gardner, D.K., Schoolcraft, W.B., 1999. Culture and transfer of human blastocysts. Curr. Opin. Obstet. Gynecol. 11 (3), 307–311. https://doi.org/10.1097/00001703-199906000-00013.

Hodes-Wertz, B., Grifo, J., Ghadir, S., Kaplan, B., Laskin, C.A., Glassner, M., Munne, S., 2012. Idiopathic recurrent miscarriage is caused mostly by aneuploid embryos. Fertil. Steril. 98 (3), 675–680. https://doi.org/10.1016/j.fertnstert.2012.j05.025.

Hynes, J.S., Forman, E.J., 2023. Transfer of the fittest: using preimplantation genetic testing for aneuploidy to select embryo(s) most likely to lead to live birth. F. S. Sci. 4 (2S), 2–6. https://doi.org/10.1016/j.xfss.2022.12.005.

Kasaven, L.S., Marcus, D., Theodorou, E., Jones, B.P., Saso, S., Naja, R., Serhal, P., Ben-Nagi, J., 2023. Systematic review and meta-analysis: does pre-implantation genetic testing for aneuploidy at the blastocyst stage improve live birth rate? J. Assist. Reprod. Genet. 40 (10), 2297–2316. https://doi.org/10.1007/s10815-023-02866-0.

Kucherov, A., Fazzari, M., Lieman, H., Ball, G.D., Doody, K., Jindal, S., 2023. PGT-A is associated with reduced cumulative live birth rate in first reported IVF stimulation cycles age ≤ 40: an analysis of 133,494 autologous cycles reported to SART CORS. J. Assist. Reprod. Genet. 40 (1), 137–149. https://doi.org/10.1007/s10815-022-02667-x.

Lee, C.I., Wu, C.H., Pai, Y.P., Chang, Y.J., Chen, C.I., Lee, T.S., Lee, M.S., 2019. Performance of preimplantation genetic testing for aneuploidy in IVF cycles for patients with advanced maternal age, repeat implantation failure, and idiopathic recurrent miscarriage. Taiwan J. Obstet. Gynecol. 58 (2), 239–243. https://doi.org/10.1016/j.tjog.2019.01.013.

Mejia, R.B., Capper, E.A., Summers, K.M., Mancuso, A.C., Sparks, A.E., Van Voorhis, B.J., 2022. Cumulative live birth rate in women aged ≤37 years after in vitro fertilization with or

without preimplantation genetic testing for aneuploidy: a Society for Assisted Reproductive Technology Clinic Outcome Reporting System retrospective analysis. F. S. Rep. 3 (3), 184–191. https://doi.org/10.1016/j.xfre.2022.05.004.

Munné, S., Kaplan, B., Frattarelli, J.L., Child, T., Nakhuda, G., Shamma, F.N., Silverberg, K., Kalista, T., Handyside, A.H., Katz-Jaffe, M., Wells, D., Gordon, T., Stock-Myer, S., Willman, S., STAR Study Group, 2019. Preimplantation genetic testing for aneuploidy versus morphology as selection criteria for single frozen-thawed embryo transfer in good-prognosis patients: a multicenter randomized clinical trial. Fertil. Steril. 112 (6), 1071–1079.e7. https://doi.org/10.1016/j.fertnstert.2019.07.1346.

Neal, S.A., Morin, S.J., Franasiak, J.M., Goodman, L.R., Juneau, C.R., Forman, E.J., Werner, M.D., Scott, Jr., R.T., 2018. Preimplantation genetic testing for aneuploidy is cost-effective, shortens treatment time, and reduces the risk of failed embryo transfer and clinical miscarriage. Fertil. Steril. 110 (5), 896–904. https://doi.org/10.1016/j.fertnstert.2018.06.021.

Ozgur, K., Berkkanoglu, M., Bulut, H., Yoruk, G.D.A., Candurmaz, N.N., Coetzee, K., 2019. Single best euploid versus single best unknown-ploidy blastocyst frozen embryo transfers: a randomized controlled trial. J. Assist. Reprod. Genet. 36 (4), 629–636. https://doi.org/10.1007/s10815-018-01399-1.

Pantou, A., Mitrakos, A., Kokkali, G., Petroutsou, K., Tounta, G., Lazaros, L., Dimopoulos, A., Sfakianoudis, K., Pantos, K., Koutsilieris, M., Mavrou, A., Kanavakis, E., Tzetis, M., 2022. The impact of preimplantation genetic testing for aneuploidies (PGT-A) on clinical outcomes in high risk patients. J. Assist. Reprod. Genet. 39 (6), 1341–1349. https://doi.org/10.1007/s10815-022-02461-9.

Rubio C., Bellver J., Rodrigo L., Castillon G., Guillen A., Vidal C., Giles J., Ferrando M., Cabanillas S., Remohi J., Pellicer A., Simon, C., 2017. In vitro fertilization with preimplantation genetic diagnosis for aneuploidies in advanced maternal age: a randomized, controlled study. 107(5):1122-1129. doi:10.1016/j.fertnstert.2017.03.011

Scheffer, J.B., Carvalho, R.F., Scheffer, B.B., Aguiar, A.P.S., Pessoa, L.P., Lozano, D.M., Fanchin, R., 2023. Correlations between clinical parameters, blastocyst morphological classification and embro euploidy. JBRA Assist. Reprod. https://doi.org/10.5935/1518-0557.20230054.

Schiewe, M.C., Whitney, J.B., Anderson, R.E., 2015. Potential risk of monochorionic dizygotic twin blastocyst formation associated with early laser zona dissection of group cultured embryos. Fertil. Steril. 103 (2), 417–421. https://doi.org/10.1016/j.fertnstert.2014.11.009.

Schoolcraft, W.B., Treff, N.R., Stevens, J.M., Ferry, K., Katz-Jaffe, M., Scott, Jr., R.T., 2011. Live birth outcome with trophectoderm biopsy, blastocyst vitrification, and single-nucleotide polymorphism microarray–based comprehensive chromosome screening in infertile patients. Fertil. Steril. 96 (3), 638–640. https://doi.org/10.1016/j.fertnstert.2011.06.049.

Surrey, E.S., Katz-Jaffe, M., Kondapalli, L.V., Gustofson, R.L., Schoolcraft, W.B., 2017. GnRH agonist administration prior to embryo transfer in freeze-all cycles of patients with endometriosis or aberrant endometrial integrin expression. Reprod. Biomed. Online 35 (2), 145–151. https://doi.org/10.1016/j.rbmo.2017.05.004.

Surrey, E.S., Katz-Jaffe, M., Surrey, R.L., Small, A.S., Gustofson, R.L., Schoolcraft, W.B., 2018. Arcuate uterus: is there an impact on in vitro fertilization outcomes after euploid embryo transfer? Fertil. Steril. 109 (4), 638–643. https://doi.org/10.1016/j.fertnstert.2017.12.001.

Tiegs, A.W., Tao, X., Zhan, Y., Whitehead, C., Kim, J., Hanson, B., Osman, E., Kim, T.J., Patounakis, G., Gutmann, J., Castelbaum, A., Seli, E., Jalas, C., Scott, Jr., R.T., 2021. A multicenter, prospective, blinded, nonselection study evaluating the predictive value of an aneuploid diagnosis using a targeted next-generation sequencing–based preimplantation genetic testing for aneuploidy assay and impact of biopsy. Fertil. Steril. 115 (3), 627–637. https://doi.org/10.1016/j.fertnstert.2020.07.052.

Werner, M., Reh, A., Grifo, J., Perle, M.A., 2012. Characteristics of chromosomal abnormalities diagnosed after spontaneous abortions in an infertile population. J. Assist. Reprod. Genet. 29 (8), 817–820. https://doi.org/10.1007/s10815-012-9781-3.

Yan, J., Qin, Y., Zhao, H, Sun, Y., Gong, F., Li, R., Sun, X., Ling, X., Li, H., Hao, C., Tan, J., Yang, J., Zhu, Y., Liu, F., Chen, D., Wei, D., Lu, J., Ni, T., Zhou, W., Wu, K., Gao, Y., Shi, Y., Lu, Y., Zhang, T., Wu, W., Ma, X., Ma, H., Fu, J., Zhang, J., Meng, Q., Zhang, H., Lego, R.S., Chen, Z.J., 2021. Live Birth with or without Preimplantation Genetic Testing for Aneuploidy. N. Engl. J. Med. 385 (22), 2047–2058. https://doi.org/10.1056/NEJMoa2103613.

Received 13 October 2023; received in revised form 23 January 2024; accepted 25 January 2024.

RBMO

ELSEVIER

ARTICLE

Clinical factors associated with unexpected poor or suboptimal response in Poseidon criteria patients

BIOGRAPHY

Alyssa Hochberg, MD, MPH, completed her obstetrics and gynaecology residency in 2021 with distinction, and is currently undertaking postgraduate training in reproductive endocrinology and infertility at McGill University, Canada. Her special research interests are obstetrical outcomes after IVF, and laboratory, obstetric and perinatal outcomes of poor and high responders.

Alyssa Hochberg[a,b,*], Michael H. Dahan[a], Hakan Yarali[c,d], Lan N. Vuong[e,f,g], Sandro C. Esteves[h]

KEY MESSAGE

In women with good ovarian reserve markers undergoing IVF, potentially modifiable clinical predictors are associated with an 'unexpected' poor or suboptimal response. This, combined with the higher-than-expected threshold values for AMH and AFC indicating such a response, may influence decisions about ovarian stimulation protocols and dose adjustments for IVF cycles.

ABSTRACT

Research question: What clinical factors are associated with 'unexpected' poor or suboptimal responses to IVF ovarian stimulation per POSEIDON's criteria, and which AMH and AFC threshold values distinguish this population?

Design: Tri-centre retrospective cohort study (2015–2017) involving first-time IVF and ICSI cycles with conventional ovarian stimulation (≥150 IU/day of FSH). Eligibility criteria included sufficient ovarian reserve markers according to POSEIDON's classification (AMH ≥1.2 ng/ml; AFC ≥5). Ovarian response categories were poor (<4 oocytes), suboptimal (4–9 oocytes) and normal (≥9 oocytes). Primary outcomes included clinical factors associated with an unexpected poor or suboptimal response to conventional ovarian stimulation using logistic regression analyses, and the threshold values of AMH and AFC predicting increased risk of such responses using ROC curves.

Results: A total of 7625 patients met the inclusion criteria: 204 (9.3%) were poor and 1998 (90.7%) were suboptimal responders. Logistic regression identified significant clinical predictors for a poor or suboptimal response, including AFC, AMH, total gonadotrophin dose, gonadotrophin type and trigger type ($P \leq 0.02$). The ROC curves indicated that AMH 2.87 ng/ml (AUC 0.740) and AFC 12 (AUC 0.826) were the threshold values predicting a poor or suboptimal response; AMH 2.17 ng/ml (AUC 0.741) and AFC 9 (AUC 0.835) predicted a poor response; and AMH 2.97 ng/ml (AUC 0.722) and AFC 12 (AUC 0.801) predicted a suboptimal response.

[a] Department of Obstetrics and Gynecology, McGill University, Montreal, Quebec, Canada.
[b] The Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.
[c] Anatolia IVF, Ankara, Turkey.
[d] Department of Obstetrics and Gynecology, Hacettepe University, Ankara, Turkey.
[e] Department of Obstetrics and Gynecology, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam.
[f] IVFMD, My Duc Hospital, Ho Chi Minh City, Vietnam.
[g] HOPE Research Center, My Duc Hospital, Ho Chi Minh City, Vietnam.
[h] ANDROFERT, Andrology and Human Reproduction Clinic, Campinas, Brazil.

KEYWORDS
Ovarian stimulation
Anti-Müllerian hormone
Antral follicle count
Poor responders
Suboptimal responders
POSEIDON criteria

*Corresponding author: E-mail address: alyssahoch@gmail.com (A. Hochberg). https://doi.org/10.1016/j.rbmo.2024.103852
1472-6483/
Declaration: SCE declares receipt of unrestricted research grants from Merck and lecture fees from Merck and MedEA; HY declares receipt of payment for lectures from Merck and Ferring; LNV receives speaker fees and conferences from Merck, Merck Sharp and Dohme (MSD) and Ferring; MHD has received unrestricted research grants from Merck Serono and Ferring.

Conclusions: The threshold values of AMH and AFC predicting 'unexpected' poor or suboptimal response were higher than expected. These findings have critical implications for tailoring IVF stimulation regimens and dosages.

INTRODUCTION

Low ovarian reserve, as indicated by abnormal ovarian reserve markers, represents a significant infertility factor (*Grisendi et al., 2019*). These markers offer diagnostic and prognostic insights, enabling clinicians to predict individual responses to ovarian stimulation during IVF cycles (*Broer et al., 2011*; *2013*). They also assist in determining the appropriate gonadotrophin stimulation dosage. Women with low ovarian reserve typically respond poorly to ovarian stimulation with exogenous gonadotrophins (*Out et al., 2000*; *Yong et al., 2003*), characterized by a low number of retrieved oocytes (*Ferraretti et al., 2011*). The effect of poor ovarian response (POR) on pregnancy rates per embryo transfer remains a matter of debate (*Morin et al., 2018*; *Shrem et al., 2022*); however, it may be associated with diminished reproductive prognosis, particularly concerning cumulative live birth rates (*Ulug et al., 2003*).

Among the most sensitive ovarian reserve markers identified to date are anti-Müllerian hormone (AMH) and antral follicle count (AFC), both readily measurable in serum and through transvaginal ultrasound, respectively (*Broekmans et al., 2010*; *La Marca and Sunkara, 2014*). Serum AMH levels and AFC in women with regular menstrual cycles decline over time, displaying a solid correlation with each other (*De Vet et al., 2002*; *Van Rooij et al., 2002*). Furthermore, the size of the recruited cohort of follicles closely relates to the remaining primordial follicular pool and AMH values (*Broekmans et al., 2010*).

Over the past decade, efforts have been made to standardize the diagnosis of POR, resulting in the development of the Bologna Criteria (*Ferraretti et al., 2011*) and the subsequent evolution of the low-prognosis patient classification known as the POSEIDON criteria (patient-oriented strategies encompassing individualized oocyte number) (*Alviggi et al., 2016*; *Esteves et al., 2016*). The POSEIDON Group (*Alviggi et al., 2016*) classification introduced four subgroups based on quantitative and qualitative parameters, such as female age and its related expected oocyte–embryo aneuploidy rate; ovarian reserve biomarkers (AFC, AMH, or both); and ovarian response, provided that a previous stimulation cycle has been carried out (*Esteves et al., 2016*; *2019*). By incorporating these parameters, the POSEIDON classification distinguishes two main categories: the 'expected' POR (groups 3 and 4) and the 'unexpected' POR (groups 1 and 2). Unexpected POR comprises groups 1 (<35 years) and 2 (≥35 years). Patients belonging to these groups have normal ovarian reserve markers (AFC ≥5, AMH ≥1.2 ng/ml, or both) but respond poorly (fewer than four oocytes retrieved; subgroups 1a and 2a) or suboptimal responses (four to nine oocytes retrieved; subgroups 1b and 2b) following conventional ovarian stimulation, i.e. using 150 IU/day or more of FSH (*Alviggi et al., 2016*; *Grisendi et al., 2019*).

The prevalence of ovarian hyporesponsiveness to gonadotrophin stimulation remains uncertain. A previous study from our research group involving 13,146 consecutive infertile women who underwent conventional ovarian stimulation, however, showed that 43% of patients met the POSEIDON criteria, with 44% and 36% falling into groups 1 and 2, respectively (*Esteves et al., 2021a*). These data suggest that about four out of 10 women with adequate pre-stimulation ovarian reserve markers experience an unexpected hypo-response to conventional ovarian stimulation.

Various explanations have been proposed to delineate the cause of hypo-response in women with ovarian reserve parameters above the POSEIDON thresholds. One prominent theory posits that these patients may harbour single nucleotide polymorphisms (SNP) in the receptor for gonadotrophins, with the most extensively studied SNP located at position 680 of the FSH receptor (*Huang et al., 2015*). Several studies have shown that patients homozygous for serine in the FSH receptor (*Alviggi et al., 2018*; *Perez Mayorga et al., 2000*), and those with a variant of the beta-subunit of the LH gene (*Alviggi et al., 2009*; *2013*; *Haahr et al., 2018*) may require increased gonadotrophin doses for ovarian stimulation. Other factors contributing to hypo-response include a low gonadotrophin starting dose (*Sunkara and Polyzos, 2018*), asynchronous follicular development and technical issues related to final oocyte maturation triggering and retrieval (*Romanski et al., 2019*).

Data on clinical factors associated specifically with an 'unexpected' poor or suboptimal response to conventional ovarian stimulation according to Poseidon's criteria are scarce. These factors may differ from those associated with an 'expected' poor or suboptimal response and from those in the entire group, and merit separate consideration. The detection of such factors may help to further delineate the mechanisms behind such 'unexpected' responses, their potential interplay and aid in appropriate planning of conventional ovarian stimulation protocols in these patients.

Moreover, in addition to exploring these clinical factors, establishing the minimum serum AMH and AFC levels associated with the retrieval of at least 10 oocytes would eliminate the need to classify responses below this threshold as 'unexpected' and better define the so-called 'normal' ovarian reserve parameters, thereby enhancing patient classification. These threshold levels would necessarily differ from those defining the 'expected' poor or suboptimal responder in the POSEIDON criteria. Therefore, the three objectives of the present study were to report the prevalence of unexpected poor and suboptimal responders among patients with adequate ovarian markers (groups 1 and 2 as per the POSEIDON criteria); to evaluate the clinical factors associated with 'unexpected poor' and 'unexpected suboptimal' responses, not previously examined; and to determine the AMH and AFC cut-off values that predict a higher risk of an unexpected poor or suboptimal response to conventional ovarian stimulation, both in combination and separately.

MATERIALS AND METHODS

Study design and participants

This retrospective cohort study included consecutive infertile women aged 22–45 years who underwent their first IVF and intracytoplasmic sperm injection cycle at three centres: ANDROFERT (Campinas, Brazil), Anatolia IVF (Ankara, Turkey), and

IVFMD at My Duc Hospital (Ho Chi Minh City, Vietnam) between 2015 and 2017.

All patients who met the following criteria were included: underwent AMH and AFC assessments within 3 months before starting their IVF cycle; had adequate pre-stimulation ovarian reserve markers according to POSEIDON's criteria (AMH ≥1.2 ng/ml and AFC ≥5); received conventional ovarian stimulation, defined as stimulation with 150 IU/day or more of FSH; and underwent oocyte retrieval. Each patient contributed data from a single IVF/ICSI cycle. Exclusion criteria comprised patients who lacked both ovarian reserve marker assessments, had discordant AMH and AFC values according to POSEIDON's criteria thresholds, underwent IVF/ICSI for purposes other than infertility, such as donor cycles, preimplantation genetic testing for monogenic diseases and fertility preservation cycles, and were treated with natural cycle IVF or mild stimulation protocols (*Nargund et al., 2007*).

Ovarian reserve assessments

Eligible patients were evaluated and treated according to the policies of their respective institutions, as previously described (*Esteves et al., 2020*). Ovarian reserve assessments were conducted during a natural menstrual cycle 1–3 months before starting stimulation, using standardized protocols. The AFC was determined during the early follicular phase via two-dimensional transvaginal ultrasonography conducted by experienced physicians from each study's centre following practical recommendations for standardized use of AFC (*Broekmans et al., 2010*). Serum AMH values were determined using the modified Beckman Coulter generation II enzyme-linked immunosorbent assay (*Craciunas et al., 2015*), as previously described (*Esteves et al., 2021b*). Both AFC and AMH assessments were conducted within the same menstrual cycle, with AMH results not available at the time of AFC assessment. Furthermore, both AFC and AMH results were accessible before selecting the ovarian stimulation regimen.

Ovarian stimulation protocols

The ovarian stimulation regimen and gonadotrophin dosage were determined based on the policies of the respective centres, considering ovarian reserve and the patient's age. Two standard ovarian stimulation protocols were used: the long gonadotrophin releasing hormone (GnRH) agonist protocol and the GnRH antagonist protocol. Patients received daily subcutaneous injections of either highly purified human menopausal gonadotrophin (HMG), recombinant-FSH, recombinant-FSH combined with HMG or recombinant-FSH combined with recombinant-LH (2:1 ratio). Daily gonadotrophin doses ranged from 150–450 IU. Ovarian response was monitored using serial transvaginal ultrasonography and oestradiol measurements, with gonadotrophin doses adjusted as required. For final oocyte maturation triggering, patients received subcutaneous injections of either recombinant-HCG (250 μg) or GnRH agonist (0.2 mg triptorelin). Oocytes were retrieved by transvaginal ultrasound-guided puncture of follicles 35–37 h after the trigger injection. Follicular fluid collected during the procedure was analysed in the IVF laboratory, and the total number of retrieved oocytes was recorded.

Data collection

Data encompassed patient demographics, infertility workup and cycle characteristics. Each patient's data was processed to establish the POSEIDON group classification per published criteria (*Alviggi et al., 2016*).

Outcomes

The primary outcomes included determining the clinical factors associated with an unexpected poor or suboptimal response to conventional ovarian stimulation, defined as the retrieval of fewer than 10 oocytes, using logistic regression analyses, and the threshold values of AMH and AFC distinguishing women at an increased risk of such responses. Additionally, we investigated the cut-off values of AMH and AFC separately, predicting an increased risk of a poor (fewer than four retrieved oocytes) or suboptimal (four to nine retrieved oocytes) response to stimulation. For this purpose, receiver operating characteristics (ROC) curves were generated using Youden's J statistic to identify the AFC and AMH values that provided the highest sensitivity and specificity (*Nahm, 2022*). This approach aligns with prior studies that used ROC curves to determine the AMH and AFC thresholds incorporated into the Poseidon criteria for predicting a poor response (*Broekmans et al., 2006*; *Broer et al., 2009*; *Grisendi et al., 2019*; *Esteves et al., 2021b*). The prevalence of unexpected poor and suboptimal responders in the dataset and the distribution of AMH and AFC values among these patient categories was also examined. Sensitivity analyses were conducted using respective ROC curves for AFC and AMH thresholds, predicting a poor or suboptimal response based on initial gonadotrophin daily doses of less than 225 IU and 225 IU or higher. These analyses evaluated whether the daily gonadotrophin dose influenced the observed results (*Lensen et al., 2018*).

Statistical analyses

Categorical data were presented by the number of cases and percentages, whereas continuous data were described using the median and interquartile range (IQR). Categorical data were subjected to Pearson's chi-squared analysis, whereas continuous data underwent Kruskal–Wallis and Wilcoxon tests as appropriate. A generalized regression approach was used incorporating the adaptive 'Least Absolute Shrinkage and Selection Operator' (Lasso) method with Akaike information criterion validation for predictor selection related to the binomial response variable 'unexpected poor or suboptimal response', both combined and for each subgroup separately (poor and suboptimal). The following predictors were simultaneously included in the model: female age, AMH, AFC, ovarian stimulation protocol, type of gonadotrophin, total gonadotrophin dose, stimulation duration, trigger type, body mass index (BMI), infertility duration and infertility factor. The study centre was considered a fixed factor. Two-way (study center versus predictor) interaction tests were conducted to investigate whether the 'study centre' had any relationship between the predictors and the response variable. Computations were carried out using JMP PRO 16 (SAS Institute, Cary, NC, USA).

Ethics

The study received approval from the Ethics Committees of the participating institutions (Instituto Investiga, Brazil, CAAE 26429219.0.0000.5599; Hacettepe University, Turkey, KA-180070; and My Duc Hospital, Vietnam, number 05/18/DD-BVMD).

RESULTS

Cohort characteristics

A total of 7625 patients were included per the criteria. An overview of cohort demographics and treatment characteristics is presented in

Supplementary Table 1. The median age of the patients was 32 years (IQR: 29–36 years), median BMI was 20.8 kg/m^2 (IQR: 19.5–22.6), median AMH was 4.25 ng/ml (IQR: 2.54–6.95), and median AFC was 15 (IQR: 10–21). Most patients (n = 7488 [97.7%]) underwent stimulation using the antagonist protocol. Most patients received recombinant FSH combined with HMG (n = 4622 [60.6%]), and most underwent ovarian stimulation with HCG (n = 6045 [79.3%]). The median number of retrieved oocytes was 13 (IQR: 9–18), and the median number of metaphase II oocytes was 9 (IQR: 6–14). The distribution of daily starting gonadotrophin doses was as follows: 150 IU (n = 1228 [16.1% of total participants]), 225 IU (n = 1598 [21.0%]), 300 IU (n = 3013 [39.5%]), and 450 IU (n = 1786 [23.4%]).

Response to stimulation

Within this cohort, 2202 patients (28.9%) exhibited a poor or suboptimal response to stimulation, with fewer than four oocytes or between four to nine retrieved oocytes, respectively (Supplementary Figure 1). Of these, 1998 patients (90.7%) were categorized as suboptimal responders, whereas 204 (9.3%) were classified as poor responders. The characteristics of these patients are presented in TABLE 1, and compares them with normal responders, defined as patients with ovarian reserve markers within the normal range as per the POSEIDON criteria and who had more than nine oocytes retrieved.

Regression analysis

A significant regression equation was established between clinical predictors and the occurrence of a poor or suboptimal response to conventional ovarian stimulation (n = 7625, R-square: 0.393) Relevant predictors included AFC (Wald chi-squared test: 411.15; P < 0.0001), AMH (Wald chi-squared test: 21.84; P < 0.0001), daily gonadotrophin dose (Wald chi-squared test: 39.64; P < 0.001), total gonadotrophin dose (Wald chi-squared test: 7.79; P = 0.005), duration of stimulation (Wald chi-squared: 16.74; P = 0.001), trigger type (Wald chi-squared test: 79.13; P < 0.0001), female age (Wald chi-squared test: 8.23; P = 0.004), type of gonadotrophin used (Wald chi-squared test: 8.71; P = 0.003), presence of any female infertility factor (Wald chi-squared test: 5.48; P = 0.02), and BMI (Wald chi-squared test: 4.96; P = 0.02) (Supplementary Table 2A). Notably, a negative relationship was found between unexpected poor or suboptimal response and AFC, AMH, ovarian stimulation with recombinant FSH (particularly when associated with LH activity, i.e., recombinant human LH or HMG), higher daily and total gonadotrophin dose, and triggering with HCG. Conversely, a positive relationship was found between unexpected poor or suboptimal response and female age, the presence of any female infertility factor, ovarian stimulation with HMG, GnRH-agonist trigger and BMI.

Subgroup logistic regression analyses using unexpected suboptimal response as the dependent variable retained all predictors mentioned above (Supplementary Table 2C). By contrast, the correspondent analysis using unexpected poor response as the dependent variable retained all the above variables except the type of gonadotrophin used for ovarian stimulation (Supplementary Table 2B).

Threshold values for response prediction

The distribution of unexpected poor or suboptimal responders according to AMH and AFC levels in the study cohort, respectively, are presented in FIGURES 1A and FIGURE 2A. The corresponding ROC curves revealed that an AMH level of 2.87 ng/ml, with an area under the curve (AUC) of 0.740 (sensitivity: 0.701; specificity: 0.662) (FIGURE 1B), and an AFC of 12, with an AUC of 0.826 (sensitivity: 0.748; specificity: 0.763) (FIGURE 2B), were the threshold values predicting a poor or suboptimal ovarian response to stimulation. The frequency of patients distributing into specific AMH and AFC levels is presented in Supplementary Figure 2.

The AMH and AFC ROC curves for predicting a poor or suboptimal response are presented in FIGURE 3. An AMH level of 2.17 ng/ml, with an AUC of 0.741 (sensitivity: 0.759, specificity: 0.617) (FIGURE 3A), and an AFC of 9, with an AUC of 0.835 (sensitivity: 0.803, specificity: 0.739) (FIGURE 3B) were identified as the threshold values predicting a poor ovarian response to stimulation. An AMH level of 2.97 ng/ml, with an AUC of 0.722 (sensitivity: 0.688, specificity: 0.651) (FIGURE 3C), and an AFC of 12, with an AUC of 0.801 (sensitivity: 0.783, specificity: 0.740) (FIGURE 3D), were determined as the threshold values predicting a suboptimal response to stimulation.

The cut-off values and area under the ROC by study centre are presented in Supplementary Table 3. In the present model, the study centre was a fixed factor. When exploring the interactions between the variable 'study centre' and predictors, overall, non-significant interactions were observed between study centres and predictors concerning the binary outcome, except for ovarian stimulation protocol (Wald chi-squared: 4.547, P = 0.03) (Supplementary Table 4).

Sensitivity analysis

In the sensitivity analyses, we focused on subjects stimulated with a daily gonadotropin starting dose of <225 IU or ≥225 IU to explore whether under-stimulation influenced the results presented. Notably, most patients received daily doses of 225 IU or more (83.9%). The ROC curves indicated that an AFC of 12 (sensitivity: 0.788; specificity: 0.723; AUC: 0.833) was the optimal threshold for predicting a poor or suboptimal ovarian response to gonadotrophin stimulation using a daily dose of less than 225 IU. Similarly, the ROC curves demonstrated that an AFC of 13 (sensitivity: 0.498; specificity: 0.886; AUC: 0.757) was the optimal threshold for predicting a poor or suboptimal ovarian response to gonadotrophin stimulation using a daily dose of 225 IU or more. The correspondent ROC curves for AMH indicated that AMH values of 2.70 ng/ml (sensitivity: 0.676; specificity: 0.777; AUC: 0.799) and 3.20 ng/ml (sensitivity: 0.610; specificity: 0.987; AUC: 0.750) were the optimal thresholds for predicting a poor or suboptimal ovarian response to gonadotrophin stimulation using daily doses of less than 225 IU and 225 IU or more, respectively.

DISCUSSION

The present study aimed to evaluate the clinical factors associated with 'unexpected poor' and 'unexpected suboptimal' responses in POSEIDON groups 1 and 2 patients, and to determine the AMH and AFC cut-off values that predict a higher risk of such a response, both in combination and separately. Our key findings can be summarized as follows: several factors were associated with decreased odds of an unexpected poor or suboptimal response, including higher AFC and AMH, conventional ovarian stimulation with recombinant FSH, a higher initial gonadotrophin dose and HCG triggering. Conversely, factors associated with increased odds of an unexpected poor or suboptimal response

TABLE 1 BASELINE CHARACTERISTICS OF UNEXPECTED POOR RESPONDERS, UNEXPECTED SUBOPTIMAL RESPONDERS, AND NORMAL RESPONDERS

Variable	Poor responders (*n* = 204 [2.7%])	Suboptimal responders (*n* = 1998 [26.2%])	Normal responders (*n* = 5423 [71.1%])	*P*-value
Age, years	36 (32–39)	34 (31–37)	32 (29–35)	<0.0001
BMI, kg/m^2	21.3 (19.8–23.6)	20.9 (19.6–22.8)	20.8 (19.5–22.5)	<0.0001
AMH	2.16 (1.45–3.42)	2.70 (1.76–4.32)	5.06 (3.15–7.88)	<0.0001
AFC	7 (6–10)	9 (7–13)	17 (13–23)	<0.0001
Duration of infertility, months	48 (24-84)	48 (24–83)	47 (24–72)	<0.0001
Infertility cause				<0.0001
Advanced maternal age	24 (11.8)	454 (22.7)	271 (5.0)	
Anatomic	0 (0.0)	13 (0.7)	56 (1.0)	
Endocrine	2 (1.0)	125 (6.3)	796 (14.7)	
Endometriosis	40 (19.6)	75 (3.8)	150 (2.8)	
Male factor	64 (31.4)	570 (28.5)	2,011 (37.1)	
Tubal factor	26 (12.7)	345 (17.3)	863 (15.9)	
Unexplained	26 (12.7)	283 (14.2)	894 (16.5)	
Other	22 (10.8)	133 (6.7)	382 (7.0)	
Ovarian stimulation protocol				
GnRH antagonist	198 (97.1)	1928 (96.5)	5322 (98.1)	<0.0001
Long GnRH agonist	6 (2.9)	70 (3.5)	101 (1.9)	
Gonadotrophins used				
HMG	1 (0.5)	13 (0.7)	31 (0.6)	<0.0001
Recombinant FSH	42 (20.6)	497 (24.9)	1948 (35.9)	
Recombinant FSH + HMG	139 (68.1)	1343 (67.2)	3140 (57.9)	
Recombinant FSH + recombinant LH	22 (10.8)	145 (7.3)	304 (5.6)	
Daily gonadotrophin dose				<0.0001
150 IU	20 (9.8)	242 (12.1)	966 (17.8)	
225 IU	23 (11.3)	265 (13.3)	1310 (24.2)	
300 IU	72 (35.3)	743 (37.2)	2198 (40.5)	
450 IU	89 (43.6)	748 (37.4)	949 (17.5)	
Total gonadotrophin dose (IU)	3000 (2400–3450)	2700 (2250–3450)	2475 (2000–3075)	<0.0001
Trigger type				
GnRH agonist trigger	12 (5.9)	94 (4.7)	1474 (27.2)	<0.0001
HCG	192 (94.1)	1904 (95.3)	3949 (72.8)	
Duration of stimulation, days	10 (8–11)	9 (8–10)	9 (8–10)	0.01
Number of oocytes retrieved	3 (2–3)	7 (6–8)	15 (12–20)	<0.0001
Number of MII oocytes	2 (1–3)	6 (4–7)	12 (10–16)	<0.0001

Continuous variables are presented as median (inter-quartile range [IQR]) and categorical variables are presented as *n* (%).

Categorical data were subjected to Pearson chi-squares analysis, whereas continuous data underwent Kruskal–Wallis and Wilcoxon tests as appropriate.

AFC, antral follicle count; AMH, anti-Müllerian hormone; BMI, body mass index; GnRH – gonadotrophin-releasing hormone; HMG, human menopausal gonadotrophin; MII, metaphase II.

included the presence of any female infertility factor, conventional ovarian stimulation with HMG, GnRH agonist triggering, increasing female age and higher BMI. The above factors hold when analysing predictors for a poor or suboptimal response separately, except for type of gonadotrophins, which was not associated with an unexpected poor response; for the combined group of unexpected and poor responders, threshold values were identified as an AMH of 2.87 ng/ml and an AFC of 12; for unexpected poor responders, specifically, the threshold values were an AMH of 2.17 ng/ml and an AFC of 9; for unexpected suboptimal responders, specifically, the threshold values were an AMH of 2.97 ng/ml and an AFC of 12. Lastly, the observed thresholds were minimally affected when considering

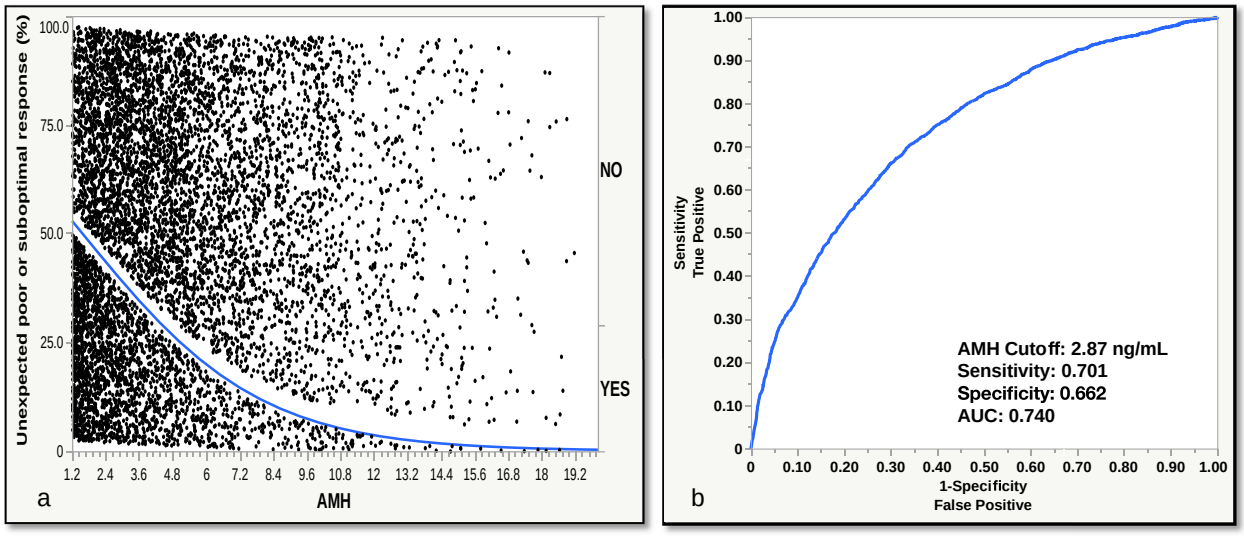


FIGURE 1 Distribution of unexpected poor or suboptimal responders according to anti-Müllerian hormone (AMH) levels and the corresponding receiver operating curve (ROC) curve. (A) Distribution of unexpected poor or suboptimal responders. The logistic plot presents the proportion of patients (each patient represented by a black dot) versus levels of AMH (ng/ml). The logistic curve (blue) displays a curve for the conditional probability, i.e. unexpected poor or suboptimal = yes/no) that depends on the continuous predictor, e.g. AMH; (b) ROC for AMH levels (ng/ml) in unexpected poor or suboptimal responders. AUC, area under the curve.

participants stimulated with less than 225 IU or 225 IU or over daily gonadotrophin doses, particularly in the case of AMH.

Notably, published data on clinical factors associated specifically with an 'unexpected' poor or suboptimal response to conventional ovarian stimulation are scarce. A cross-sectional study (*Noli et al., 2023*) examined the association between adherence to a Mediterranean diet (assessed through a Mediterranean Diet Score) and the risk of unexpected poor response (three or fewer retrieved oocytes). They included a total of 296 women with adequate ovarian reserve (defined as an AFC between 10–22 or an AMH between 2–5 ng/ml), of whom 47 were poor responders. This study demonstrated that low adherence to a Mediterranean diet could be a risk factor for unexpected poor ovarian response. However 'unexpected' poor response was not based on the agreed-upon ovarian reserve thresholds proposed by Poseidon. Another recent study (*Mehdizadeh et al., 2023*) examined novel Bone morphogenetic protein15 (BMP15) and growth differentiation factor 9 (GDF9) variants in 'unexpected' poor or

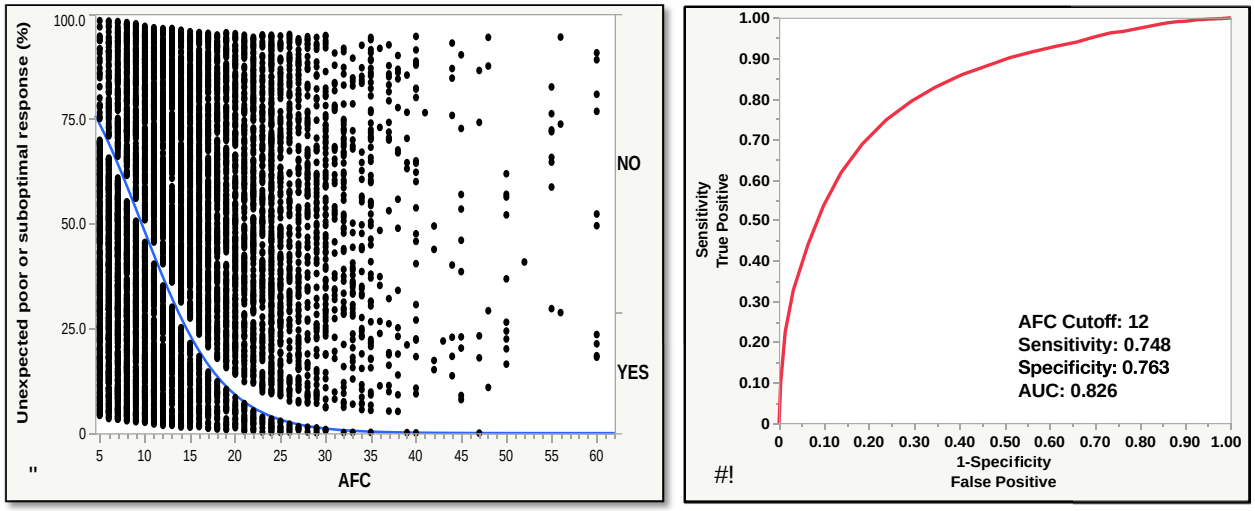


FIGURE 2 Distribution of unexpected poor or suboptimal responders according to antral follicle count (AFC) and the corresponding receiver operating characteristic (ROC) curve. (A) Distribution of unexpected poor or suboptimal responders. The logistic plot presents the proportion of patients (each patient represented by a black dot) versus levels of AFC. The logistic curve (blue) displays a curve for the conditional probability, i.e. unexpected poor or suboptimal = yes/no) that depends on the continuous predictor, e.g. AFC; (B) ROC for AFC in unexpected poor or suboptimal responders. AUC, area under the curve.

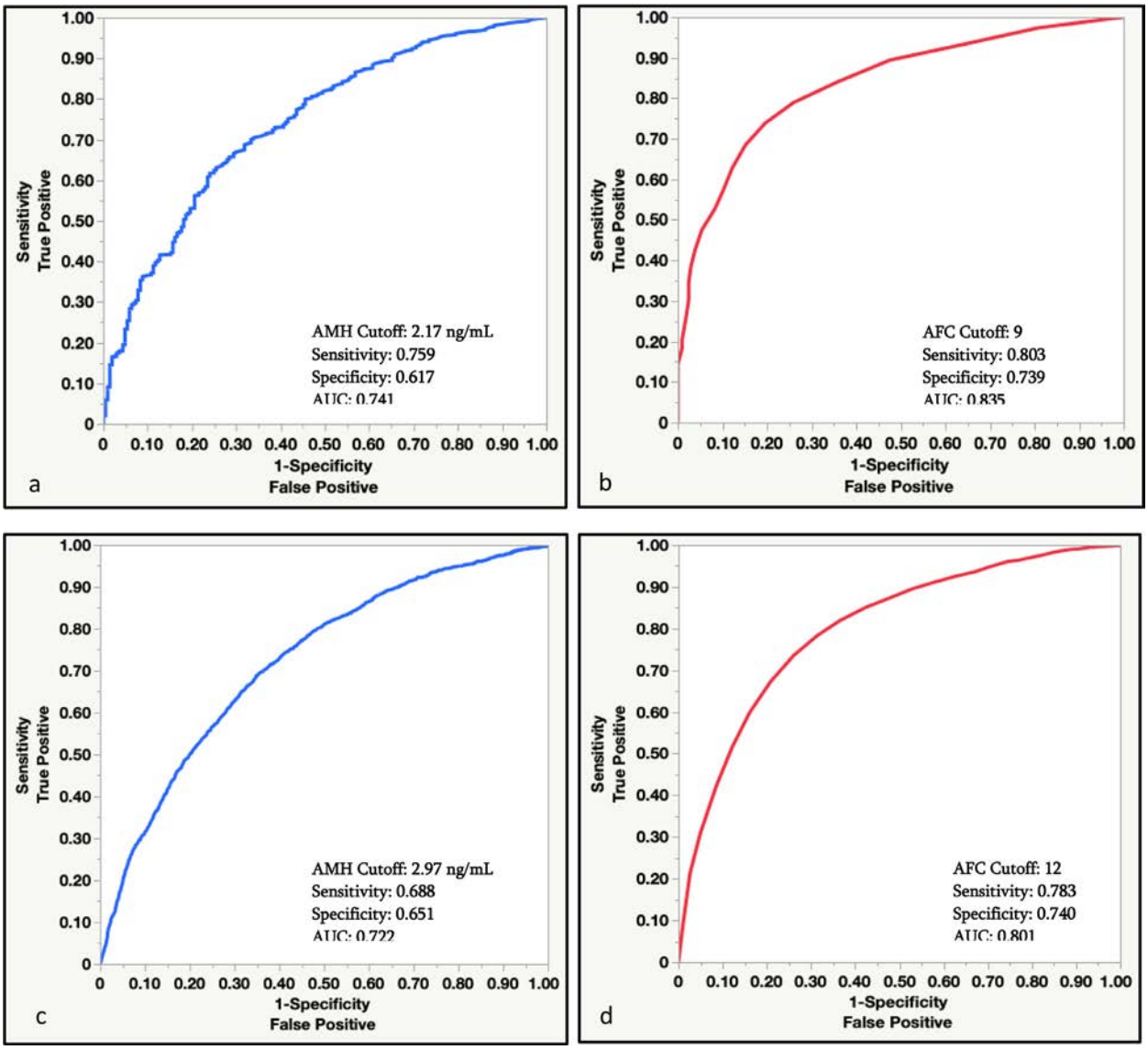


FIGURE 3 Receiver operating characteristic (ROC) curves for (A) anti-Müllerian hormone (AMH) levels and (B) antral follicle count (AFC) in unexpected poor responders; (C) AMH levels and (D) AFC in unexpected suboptimal responders. AUC, area under the curve.

suboptimal ovarian response (defined as nine or fewer retrieved oocytes, without distinguishing between poor and suboptimal responses), compared with normal responders. They found that missense variants were found exclusively in poor or suboptimal responders; and that the mutant allele (T) in a GDF9 polymorphism (C447T) was found more frequently in normal compared with poor or suboptimal responders. However 'unexpected' poor response as an AMH at and above 1.27 ng/ml (slightly above the Poseidon thresholds) and without taking AFC into account when defining 'normal' ovarian reserve. Unlike these two previously published studies (*Noli et al., 2023*; *Mehdizadeh et al., 2023*) in the present study, myriad baseline and treatment clinical predictors relevant to cycle planning were examined, which had not previously been investigated from the perspective of 'unexpected' poor or suboptimal response risk, which may potentially be modifiable.

The present study revealed that triggering final follicle maturation with HCG decreased the likelihood of experiencing an unexpectedly poor or suboptimal response to ovarian stimulation. Previous studies (*Chen et al., 2012*; *Kummer et al., 2013*; *Meyer et al., 2015*) have suggested that a small subgroup of patients who did not respond adequately to a GnRH-agonist trigger, as indicated by an LH surge of less than 15 mIU/ml, were more likely to have significantly lower rates of oocyte retrieval and overall cycle success. In line with earlier research (*Hompes et al., 2008*; *Lehert et al., 2010*; *Chen et al., 2012*; *Devroey et al., 2012*; *Kummer et al., 2013*; *Meyer et al., 2015*; *Bernstein et al., 2019*), the present study also found that several other factors were associated with an increased risk of experiencing an unexpectedly poor or suboptimal response to stimulation. These factors included the presence of any female infertility condition, ovarian stimulation with HMG, GnRH agonist trigger, higher female age, and an elevated BMI. Regarding the association between ovarian stimulation with HMG and increased odds of a poor or

suboptimal response, our finding accords with previous studies demonstrating a higher oocyte yield when using recombinant FSH compared with highly purified HMG, both when using the long GnRH agonist protocol (*Hompes et al., 2008*), and when analysed in a subsequent meta-analysis of 16 randomized controlled trials (*Lehert et al., 2010*). A proposed explanation for this phenomenon (*Bosch et al., 2008*) has been that LH activity, provided by HCG in HMG preparations, may induce atresia of some of the follicles during the follicular phase, as also observed when comparing HMG and recombinant FSH in ovulation induction in World Health Organization Group II anovulatory infertility patients (*Platteau et al., 2006*), where fewer intermediate follicles were present in patients receiving highly purified HMG. This explanation regarding LH activity, however, does not account for the observation in our study that recombinant FSH, even supplemented with recombinant LH, was associated with a reduced likelihood of a poor or suboptimal response. It is possible that the unique properties of recombinant FSH, whether used alone, or in combination with recombinant LH, make it more effective for follicular recruitment compared with urinary FSH or HMG in patients at risk of a poor or suboptimal ovarian response. Nevertheless, this hypothesis would require further investigation through pharmacologic studies.

The association between the presence of female infertility factors and the risk of a poor or suboptimal response aligns with recent findings that demonstrated an elevated risk of an unexpected poor response in women with endometriosis compared with controls undergoing their initial IVF cycle (*Invernici et al., 2022*). Additionally, it corresponds with the well-established link between advanced maternal age and IVF outcomes (*Bernstein et al., 2019*).

Our results of the effect of increasing BMI on reproductive outcomes are consistent with a previous study (*Dornelles et al., 2022*), which found fewer retrieved oocytes and mature oocytes in women with a BMI of 25 kg/m^2 or higher compared with those women with a BMI below 25 kg/m^2. Furthermore, in line with previous studies (*Behre et al., 2005*; *Chen et al., 2012*; *Kummer et al., 2013*; *Meyer et al., 2015*; *Drakopoulos et al., 2018*; *Bernstein et al., 2019*; *Invernici et al., 2022*), our findings indicate that AFC and AMH, the gonadotrophin starting dose, trigger type, advanced female age and presence of female infertility factor were associated with an unexpected poor response to stimulation.

Regarding the AMH and AFC thresholds we detected, our findings underscore that women with truly normal ovarian reserve seem to be characterized by an AFC of 13 or greater and a serum AMH value of at least 2.88 ng/ml when assessed using the Beckman Coulter assay. Notably, limited data are available for identifying women with adequate ovarian reserve parameters who may be at risk of experiencing a poor or suboptimal response to conventional ovarian stimulation. Similarly, information on AMH and AFC cut-off values that predict such a response is scarce. A retrospective study (*Wang et al., 2021*), which investigated cumulative pregnancy rates and live birth rates in patients undergoing IVF/ICSI using the GnRH agonist long protocol and daily doses of recombinant FSH ranging from 150 to 225 IU, found no significant difference in mean AMH levels between suboptimal and normal responders. In the present study, suboptimal response was defined as serum oestradiol levels falling below 658.8 pmol/l on days 6–8 of stimulation, along with ultrasound evidence of at least six follicles ranging from 6–10 mm in size and no follicle with a mean diameter exceeding 10 mm. The study (*Wang et al., 2021*) included women aged 20–40 years with an AFC of 5–7 or more, AMH levels of 0.5–1.1 ng/ml or higher and basal day-3 FSH levels under 10 IU/l. The investigators identified age, BMI and high basal FSH levels as independent risk factors for a suboptimal response. Similar to our study, they observed that a higher initial gonadotrophin dose was an independent protective factor against a suboptimal response. Unlike the present study, this previous research did not classify poor and suboptimal responders using the POSEIDON criteria or use ROC curves to establish cut-off values for predicting a suboptimal response.

Another retrospective study (*Alvaro Mercadal et al., 2018*) involving 735 women undergoing IVF with the GnRH antagonist protocol found significantly lower mean AMH and AFC values among suboptimal (those with four to nine oocytes retrieved) than optimal (those with 10–15 oocytes retrieved) responders. Specifically, the study reported mean AMH levels of 1.15 ± 1.31 ng/ml for suboptimal responders and 2.49 ± 2.41 ng/ml for optimal responders. Similarly, the mean AFC was 9.7 ± 5.4 for suboptimal responders and 15.2 ± 6.8 for optimal responders. These differences remained significant even after adjusting for confounding variables, including age, and were associated with variations in the cumulative pregnancy rates between the two groups. Unlike the present study, this research did not limit its inclusion criteria to women with adequate ovarian reserve parameters as defined by POSEIDON's criteria, nor did it explore the cut-off values of AMH and AFC that might predict a suboptimal response.

In line with our findings, a previous study conducted by our research team (*Esteves et al., 2021b*), involving a retrospective analysis of 9484 consecutive infertility patients, revealed that the respective cut-off values for AFC and AMH distinguishing patients at risk of suboptimal response or not were 12 and 2.97 ng/ml. The sensitivities of these cut-offs were determined to be 0.74 and 0.69, with specificities of 0.76 and 0.66. The AUC values for these cut-offs were 0.81 and 0.80, indicating their effectiveness. This research encompassed a diverse range of infertile patients who met various POSEIDON group criteria, including patients not classified within the POSEIDON groups.

In contrast, our current investigation focused on a sizable cohort of patients who specifically met the criteria outlined in POSEIDON groups 1 and 2. Remarkably, our results aligned with the previous research (*Esteves et al., 2021b*) on AMH and AFC threshold values predicting a suboptimal response to ovarian stimulation. These values had similar thresholds, AUCs, sensitivities and specificities, thereby reinforcing the robustness of our findings. The present study is novel compared with our previously published study in several respects. First, we included unexpected poor responders as a separate category with distinct AMH and AFC cut-offs, as well as a combined category for both suboptimal and poor response. To the best of our knowledge, the present study is the first to explore and define the threshold values of AMH and AFC that predict an unexpectedly poor or suboptimal response when considering both conditions together and separately. This unique aspect of our study contributes to the existing body of knowledge in this field, as it has not been previously examined. Second, we excluded patients with

discordant ovarian reserve markers as they may exhibit different AMH and AFC thresholds predicting a poor or suboptimal response, potentially representing a different patient population. Thereby, our cohort consisted exclusively of patients with concordant ovarian biomarkers, resulting in a more homogenous population of patients with a 'true' unexpected poor or suboptimal response to conventional ovarian stimulation. Third, our study is unique in that it examines clinical predictors of an 'unexpected' poor or suboptimal response, factors not previously explored in this subgroup. Future studies will need to investigate whether adjusting treatment for these patients according to the associated factors we identified can improve oocyte yield. This would help isolate the effect of each specific factor as well as the effect of FSH/LH receptor polymorphism.

Notably, the original cut-offs specified in the POSEIDON criteria were derived from data in the general population. As outlined in the criteria themselves, however, these cut-offs did not accurately apply to the group of 'unexpected' poor or suboptimal responders. Our aim was to investigate these thresholds specifically within the subset of individuals classified as 'unexpected' poor or suboptimal responders, who, according to the POSEIDON criteria, possess 'adequate' ovarian reserve parameters. Although we did not analyse these thresholds in the broader population, encompassing all response types, we believe these thresholds remain pertinent for predicting an increased risk of being categorized as an 'unexpected' poor or suboptimal responder.

The primary strength of our study lies in its substantial sample size, allowing for the detection of clinical predictors of an 'unexpected' poor or suboptimal response, factors which have not been previously evaluated in this specific subgroup. These add to current knowledge about possible causes and potentially effective treatments in these challenging patients. This large sample size also allowed us to construct ROC curves with relatively accurate AMH and AFC thresholds. Additionally, using a multicentre database enabled us to enhance the generalizability across three different populations: South American (predominantly Brazilian), Southeast Asian, and Turkish European, including Middle Eastern populations. To the best of our knowledge, this study represents the most extensive investigation assessing these threshold levels for predicting an unexpected poor or suboptimal response, both in combination and individually, among POSEIDON groups 1 and 2 patients. Lastly, our ability to detect readily measurable ovarian reserve markers and a myriad of clinical factors before starting stimulation may assist in guiding and defining appropriate management before ovarian stimulation.

The present study has several limitations. First, its retrospective nature means that there may be undetected biases that could influence the results. For example, data on cycle cancellation rates were not available. As we included only patients with 'normal' ovarian reserve markers per POSEIDON, however, the number of cancelled cycles would be expected to be low and, therefore, not significantly affecting our results. Second, the measurement of AFC involved various operators using different machines. Although inter-observer variability of AFC has been reported to be generally low (*Subirá et al., 2017*), we cannot completely rule out the possibility that differences in technique and reporting may have influenced the accuracy of AFC measurements. Additionally, AMH measurements were conducted in different laboratories, although they used the same manual assay. Although automated assays have been associated with lower inter-laboratory variability than manual assays (*Gassner and Jung, 2014*), the latter was the only available method during the study period. Furthermore, one must acknowledge the limitations of establishing thresholds for continuous variables by ROC curves, as studies using distinct populations may find slightly different thresholds than those we detected, limiting the generalizability of these specific thresholds. Nonetheless, these thresholds, computed similarly to those proposed by the POSEIDON criteria, may still serve as guidance when assessing each patient's clinical characteristics before the start of ovarian stimulation.

Furthermore, it is essential to note that ovarian response may vary among different populations because of genetic variations in FSH receptors. Unfortunately, the present study did not account for this factor. Although we do not know the exact mechanism by which the clinical factors we detected play a role or interact with the gonadotrophin receptor polymorphism theory, it may be reasonable to assume that the cause behind a suboptimal response is multi-factorial, including genetic and clinical factors, that may interplay. In future research, it is essential to consider this genetic variation when analysing ovarian responses. Furthermore, studies should also aim to investigate the distinct effects of various treatments, such as trigger type and FSH dose. This approach can provide deeper insights into the role of these clinical factors compared with the influence of 'pure' genetic variations in this patient population. Ideally, the repetition of this study in a large prospective database would provide more robust insights into these ovarian reserve markers and their predictive value for ovarian stimulation outcomes.

In conclusion, the present study has identified various baseline and treatment characteristics influencing the risk of experiencing an 'unexpected' poor or suboptimal response to ovarian stimulation, aiding in understanding the intricacies of this unique subgroup. Additionally, we detected specific threshold values for AMH and AFC that can predict an unexpectedly poor or suboptimal response, both together and separately. These values are AMH 2.87 ng/ml or lower and an AFC 12 or less for the combined group, AMH 2.17 ng/ml or lower and AFC 9 or less for unexpected poor responders, and AMH 2.97 ng/ml or lower and an AFC 12 or less for suboptimal responders. When clinicians encounter patients with ovarian reserve parameters at or below these identified values, especially in the context of the clinical predictors mentioned above, they should consider the potential for a poor or suboptimal response to stimulation and tailor treatment, including FSH dosage, gonadotrophin regimen and trigger type. We believe that using these markers and predictors to proactively identify such patients is of utmost importance, as it enables more personalized treatment approaches, maximizes the chances of obtaining an adequate number of oocytes and ultimately enhances the reproductive potential of these individuals.

DATA AVAILABILITY

Data will be made available on request.

ACKNOWLEDGEMENTS

We thank Luciana Abreu and Professor José Carvalho for their help with data compilation and statistical analyses.

AUTHORS' ROLES

AH, MHD, and SCE designed the study and helped with data acquisition, analysis and interpretation; HY and LV participated in data acquisition, data interpretation and article revision for critical intellectual content; all authors contributed intellectually to the writing or revision of the manuscript, approved the final version and are accountable for all aspects of the work.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.rbmo.2024.103852.

REFERENCES

Alvaro Mercadal, B., Rodríguez, I., Arroyo, G., Martínez, F., Barri, P.N., Coroleu, B., 2018. Characterization of a suboptimal IVF population and clinical outcome after two IVF cycles. Gynecol Endocrinol 34, 125–128. https://doi.org/10.1080/09513590.2017.1369515.

Alviggi, C., Andersen, C.Y., Buehler, K., Conforti, A., De Placido, G., Esteves, S.C., Fischer, R., Galliano, D., Polyzos, N.P., Sunkara, S.K., Ubaldi, F.M., Humaidan, P., 2016. A new more detailed stratification of low responders to ovarian stimulation: from a poor ovarian response to a low prognosis concept. Fertil Steril 105, 1452–1453. https://doi.org/10.1016/J.FERTNSTERT.2016.02.005.

Alviggi, C., Conforti, A., Esteves, S.C., Andersen, C.Y., Bosch, E., Bühler, K., Ferraretti, A.P., De Placido, G., Mollo, A., Fischer, R., Humaidan, P., 2018. Recombinant luteinizing hormone supplementation in assisted reproductive technology: a systematic review. Fertil Steril 109, 644–664. https://doi.org/10.1016/J.FERTNSTERT.2018.01.003.

Alviggi, C., Humaidan, P., Howles, C.M., Tredway, D., Hillier, S.G., 2009. Biological versus chronological ovarian age: implications for assisted reproductive technology. Reprod Biol Endocrinol 7, 101. https://doi.org/10.1186/1477-7827-7-101.

Alviggi, C., Pettersson, K., Longobardi, S., Andersen, C.Y., Conforti, A., De Rosa, P., Clarizia, R., Strina, I., Mollo, A., De Placido, G., Humaidan, P., 2013. A common polymorphic allele of the LH beta-subunit gene is associated with higher exogenous FSH consumption during controlled ovarian stimulation for assisted reproductive technology. Reprod Biol Endocrinol 11. https://doi.org/10.1186/1477-7827-11-51.

Behre, H.M., Greb, R.R., Mempel, A., Sonntag, B., Kiesel, L., Kaltwaßer, P., Seliger, E., Röpke, F., Gromoll, J., Nieschlag, E., Simoni, M., 2005. Significance of a common single nucleotide polymorphism in exon 10 of the follicle-stimulating hormone (FSH) receptor gene for the ovarian response to FSH: a pharmacogenetic approach to controlled ovarian hyperstimulation. Pharmacogenet Genomics 15, 451–456. https://doi.org/10.1097/01.FPC.0000167330.92786.5E.

Bernstein, L.R., Nawroth, F., Bedoschi, G.M., Cimadomo, D., Ubaldi, F.M., Vaiarelli, A., Fabozzi, G., Venturella, R., Maggiulli, R., Mazzilli, R., Ferrero, S., Palagiano, A., Rienzi, L., 2019. Advanced Maternal Age in IVF: Still a Challenge? The Present and the Future of Its Treatment. Frontiers in Endocrinology | www.frontiersin.org 10, 94. https://doi.org/10.3389/fendo.2019.00094.

Bosch, E., Vidal, C., Labarta, E., Simon, C., Remohi, J., Pellicer, A., 2008. Highly purified hMG versus recombinant FSH in ovarian hyperstimulation with GnRH antagonists—a randomized study. Hum Reprod 23, 2346–2351. https://doi.org/10.1093/HUMREP/DEN220.

Broekmans, F.J., Kwee, J., Hendriks, D.J., Mol, B.W., Lambalk, C.B., 2006. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update 12, 685–718. https://doi.org/10.1093/HUMUPD/DML034.

Broekmans, F.J.M., De Ziegler, D., Howles, C.M., Gougeon, A., Trew, G., Olivennes, F., 2010. The antral follicle count: practical recommendations for better standardization. Fertil Steril 94, 1044–1051. https://doi.org/10.1016/J.FERTNSTERT.2009.04.040.

Broer, S.L., Dólleman, M., Opmeer, B.C., Fauser, B.C., Mol, B.W., Broekmans, F.J.M., 2011. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis. Hum Reprod Update 17, 46–54. https://doi.org/10.1093/HUMUPD/DMQ034.

Broer, S.L., Mol, B.W.J., Hendriks, D., Broekmans, F.J.M., 2009. The role of antimullerian hormone in prediction of outcome after IVF: comparison with the antral follicle count. Fertil Steril 91, 705–714. https://doi.org/10.1016/J.FERTNSTERT.2007.12.013.

Broer, S.L., van Disseldorp, J., Broeze, K.A., Dolleman, M., Opmeer, B.C., Bossuyt, P., Eijkemans, M.J.C., Mol, B.W.J., Broekmans, F.J.M., Anderson, R.A., Ashrafi, M., Bancsi, L., Caroppo, E., Copperman, A., Ebner, T., Geva, M.E., Erdem, M., Greenblatt, E.M., Jayaprakasan, K., Fenning, R., Klinkert, E.R., Kwee, J., Lambalk, C.B., La Marca, A., McIlveen, M., Merce, L.T., Muttukrishna, S., Nelson, S.M., Ng, H.Y., Popovic-Todorovic, B., Smeenk, J.M.J., Tomás, C., Van der Linden, P.J.Q., van Rooij, I.A., Vladimirov, I.K., 2013. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum Reprod Update 19, 26–36. https://doi.org/10.1093/HUMUPD/DMS041.

Chen, S.L., Ye, D.S., Chen, X., Yang, X.H., Zheng, H.Y., Tang, Y., He, Y.X., Guo, W., 2012. Circulating luteinizing hormone level after triggering oocyte maturation with GnRH agonist may predict oocyte yield in flexible GnRH antagonist protocol. Hum Reprod 27, 1351–1356. https://doi.org/10.1093/HUMREP/DES049.

Craciunas, L., Roberts, S.A., Yates, A.P., Smith, A., Fitzgerald, C., Pemberton, P.W., 2015. Modification of the Beckman-Coulter second-generation enzyme-linked immunosorbent assay protocol improves the reliability of serum antimüllerian hormone measurement. Fertil Steril 103,. https://doi.org/10.1016/J.FERTNSTERT.2014.10.052 554-559.e1.

De Vet, A., Laven, J.S.E., De Jong, F.H., Themmen, A.P.N., Fauser, B.C.J.M., 2002. Antimüllerian hormone serum levels: A putative marker for ovarian aging. Fertil Steril 77, 357–362. https://doi.org/10.1016/S0015-0282(01)02993-4.

Devroey, P., Pellicer, A., Nyboe Andersen, A., Arce, J.C., 2012. A randomized assessor-blind trial comparing highly purified hMG and recombinant FSH in a GnRH antagonist cycle with compulsory single-blastocyst transfer. Fertil Steril 97, 561–571. https://doi.org/10.1016/J.FERTNSTERT.2011.12.016.

Dornelles, V.C., Hentschke, M.R., Badalotti, M., Telöken, I.B., Trindade, V.D., Cunegatto, B., de Vasconcelos, N.F., da Costa, B.E.P., Petracco, A., Padoin, A.V., 2022. The impact of body mass index on laboratory, clinical outcomes and treatment costs in assisted reproduction: a retrospective cohort study. BMC Womens Health 22. https://doi.org/10.1186/S12905-022-02036-X.

Drakopoulos, P., Santos-Ribeiro, S., Bosch, E., Garcia-Velasco, J., Blockeel, C., Romito, A., Tournaye, H., Polyzos, N.P., 2018. The Effect of Dose Adjustments in a Subsequent Cycle of Women With Suboptimal Response Following Conventional Ovarian Stimulation. Front

Endocrinol (Lausanne) 9. https://doi.org/10.3389/FENDO.2018.00361.
Esteves, S.C., Carvalho, J.F., Bento, F.C., Santos, J., 2019. A Novel Predictive Model to Estimate the Number of Mature Oocytes Required for Obtaining at Least One Euploid Blastocyst for Transfer in Couples Undergoing in vitro Fertilization/Intracytoplasmic Sperm Injection: The ART Calculator. Front Endocrinol (Lausanne) 10. https://doi.org/10.3389/FENDO.2019.00099.
Esteves, S.C., Humaidan, P., Alviggi, C., Fischer, R., 2016. The novel POSEIDON stratification of "Low prognosis patients in Assisted Reproductive Technology" and its proposed marker of successful outcome. F1000Res 5. https://doi.org/10.12688/F1000RESEARCH.10382.1/DOI.
Esteves, S.C., Yarali, H., Ubaldi, F.M., Carvalho, J.F., Bento, F.C., Vaiarelli, A., Cimadomo, D., Özbek, İ.Y., Polat, M., Bozdag, G., Rienzi, L., Alviggi, C., 2020. Validation of ART Calculator for Predicting the Number of Metaphase II Oocytes Required for Obtaining at Least One Euploid Blastocyst for Transfer in Couples Undergoing in vitro Fertilization/Intracytoplasmic Sperm Injection. Front Endocrinol (Lausanne) 10. https://doi.org/10.3389/FENDO.2019.00917.
Esteves, S.C., Yarali, H., Vuong, L.N., Carvalho, J.F., Özbek, İ.Y., Polat, M., Le, H.L., Pham, T.D., Ho, T.M., 2021a. Low Prognosis by the POSEIDON Criteria in Women Undergoing Assisted Reproductive Technology: A Multicenter and Multinational Prevalence Study of Over 13,000 Patients. Front Endocrinol (Lausanne) 12. https://doi.org/10.3389/FENDO.2021.630550.
Esteves, S.C., Yarali, H., Vuong, L.N., Carvalho, J.F., Özbek, İ.Y., Polat, M., Le, H.L., Pham, T.D., Ho, T.M., 2021b. Antral follicle count and anti-Müllerian hormone to classify low-prognosis women under the POSEIDON criteria: a classification agreement study of over 9000 patients. Hum Reprod 36, 1530–1541. https://doi.org/10.1093/HUMREP/DEAB056.
Ferraretti, A.P., La Marca, A., Fauser, B.C.J.M., Tarlatzis, B., Nargund, G., Gianaroli, L., 2011. ESHRE consensus on the definition of "poor response" to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum Reprod 26, 1616–1624. https://doi.org/10.1093/HUMREP/DER092.
Gassner, D., Jung, R., 2014. First fully automated immunoassay for anti-Müllerian hormone. Clin Chem Lab Med 52, 1143–1152. https://doi.org/10.1515/CCLM-2014-0022.
Grisendi, V., Mastellari, E., La Marca, A., 2019. Ovarian Reserve Markers to Identify Poor Responders in the Context of Poseidon Classification. Front Endocrinol (Lausanne) 10. https://doi.org/10.3389/FENDO.2019.00281.
Haahr, T., Esteves, S.C., Humaidan, P., 2018. Individualized controlled ovarian stimulation in expected poor-responders: an update. Reprod Biol Endocrinol 16. https://doi.org/10.1186/S12958-018-0342-1.
Hompes, P.G.A., Broekmans, F.J., Hoozemans, D.A., Schats, R., 2008. Effectiveness of highly purified human menopausal gonadotropin vs. recombinant follicle-stimulating hormone in first-cycle in vitro fertilization-intracytoplasmic sperm injection patients. Fertil Steril 89, 1685–1693. https://doi.org/10.1016/J.FERTNSTERT.2007.05.039.
Huang, X., Li, L., Hong, L., Zhou, W., Shi, H., Zhang, H., Zhang, Z., Sun, X., Du, J., 2015. The Ser680Asn polymorphism in the follicle-stimulating hormone receptor gene is associated with the ovarian response in controlled ovarian hyperstimulation. Clin Endocrinol (Oxf) 82, 577–583. https://doi.org/10.1111/CEN.12573.
Invernici, D., Reschini, M., Benaglia, L., Somigliana, E., Galati, G., La Vecchia, I., Vigano, P., Vercellini, P., 2022. The impact of endometriosis on IVF efficacy: qualitative and quantitative assessment of ovarian response and embryo development. Reprod Biomed Online 45, 275–281. https://doi.org/10.1016/J.RBMO.2022.04.010.
Kummer, N.E., Feinn, R.S., Griffin, D.W., Nulsen, J.C., Benadiva, C.A., Engmann, L.L., 2013. Predicting successful induction of oocyte maturation after gonadotropin-releasing hormone agonist (GnRHa) trigger. Hum Reprod 28, 152–159. https://doi.org/10.1093/HUMREP/DES361.
La Marca, A., Sunkara, S.K., 2014. Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: from theory to practice. Hum Reprod Update 20, 124–140. https://doi.org/10.1093/HUMUPD/DMT037.
Lehert, P., Schertz, J.C., Ezcurra, D., 2010. Recombinant human follicle-stimulating hormone produces more oocytes with a lower total dose per cycle in assisted reproductive technologies compared with highly purified human menopausal gonadotrophin: a meta-analysis. Reprod Biol Endocrinol 8. https://doi.org/10.1186/1477-7827-8-112.
Lensen, S.F., Wilkinson, J., Leijdekkers, J.A., La Marca, A., Mol, B.W.J., Marjoribanks, J., Torrance, H., Broekmans, F.J., 2018. Individualised gonadotropin dose selection using markers of ovarian reserve for women undergoing in vitro fertilisation plus intracytoplasmic sperm injection (IVF/ICSI). Cochrane Database Syst Rev 2. https://doi.org/10.1002/14651858.CD012693.PUB2.
Mehdizadeh, A., Soleimani, M., Amjadi, F., Sene, A.A., Sheikhha, M.H., Dehghani, A., Ashourzadeh, S., Aali, B.S., Dabiri, S., Zandieh, Z., 2023. Implication of Novel BMP15 and GDF9 Variants in Unexpected Poor Ovarian Response. Reprod Sci. https://doi.org/10.1007/S43032-023-01370-1.
Meyer, L., Murphy, L.A., Gumer, A., Reichman, D.E., Rosenwaks, Z., Cholst, I.N., 2015. Risk factors for a suboptimal response to gonadotropin-releasing hormone agonist trigger during in vitro fertilization cycles. Fertil Steril 104, 637–642. https://doi.org/10.1016/J.FERTNSTERT.2015.06.011.
Morin, S.J., Patounakis, G., Juneau, C.R., Neal, S.A., Scott, R.T., Seli, E., 2018. Diminished ovarian reserve and poor response to stimulation in patients. Hum Reprod 33, 1489–1498. https://doi.org/10.1093/HUMREP/DEY238.
Nahm, F.S., 2022. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol 75, 25. https://doi.org/10.4097/KJA.21209.
Nargund, G., Fauser, B.C.J.M., Macklon, N.S., Ombelet, W., Nygren, K., Frydman, R., 2007. The ISMAAR proposal on terminology for ovarian stimulation for IVF. Hum Reprod 22, 2801–2804. https://doi.org/10.1093/HUMREP/DEM285.
Noli, S.A., Ferrari, S., Ricci, E., Reschini, M., Cipriani, S., Dallagiovanna, C., Parazzini, F., Somigliana, E., 2023. Adherence to the Mediterranean diet and the risk of unexpected poor response to ovarian stimulation in IVF cycles. Reprod Biomed Online 47, 77–83. https://doi.org/10.1016/J.RBMO.2023.03.011.
Out, H.J., Braat, D.D.M., Lintsen, B.M.E., Gurgan, T., Bukulmez, O., Gökmen, O., Keles, G., Caballero, P., González, J.M., Fábregues, F., Balasch, J., Roulier, R., 2000. Increasing the daily dose of recombinant follicle stimulating hormone (Puregon) does not compensate for the age-related decline in retrievable oocytes after ovarian stimulation. Hum Reprod 15, 29–35. https://doi.org/10.1093/HUMREP/15.1.29.
Perez Mayorga, M., Gromoll, J., Behre, H.M., Gassner, C., Nieschlag, E., Simoni, M., 2000. Ovarian response to follicle-stimulating hormone (FSH) stimulation depends on the FSH receptor genotype. J Clin Endocrinol Metab 85, 3365–3369. https://doi.org/10.1210/JCEM.85.9.6789.
Platteau, P., Andersen, A.N., Balen, A., Devroey, P., Sørensen, P., Helmgaard, L., Arce, J.C., 2006. Similar ovulation rates, but different follicular development with highly purified menotrophin compared with recombinant FSH in WHO Group II anovulatory infertility: a randomized controlled study. Hum Reprod 21, 1798–1804. https://doi.org/10.1093/HUMREP/DEL085.
Romanski, P.A., Farland, L.V., Tsen, L.C., Ginsburg, E.S., Lewis, E.I., 2019. Effect of class III and class IV obesity on oocyte retrieval complications and outcomes. Fertil Steril 111,. https://doi.org/10.1016/J.FERTNSTERT.2018.10.015 294-301.e1.
Shrem, G., Salmon-Divon, M., Mahfoudh, A.M., Balayla, J., Volodarsky-Perel, A., Henderson, S., Zeadna, A., Son, W.Y., Steiner, N., Dahan, M.H., 2022. Influence of Maternal Age and Ovarian Reserve on the Decision to Continue or to Cancel IVF Cycles in Patients with One or Two Large Follicles: a Dual Effect. Reprod Sci 29, 291–300. https://doi.org/10.1007/S43032-021-00649-5.
Subirá, J., Alberola-Rubio, J., Núñez, M.J., Escrivá, A.M., Pellicer, A., Montañana, V., Díaz-García, C., 2017. Inter-cycle and inter-observer variability of the antral follicle count in routine clinical practice. Gynecol Endocrinol 33, 515–518. https://doi.org/10.1080/09513590.2017.1291614.
Sunkara, S.K., Polyzos, N.P., 2018. OPTIMIST trial: optimistic evidence? Hum Reprod 33, 983–984. https://doi.org/10.1093/HUMREP/DEY062.
Ulug, U., Ben-Shlomo, I., Turan, E., Erden, H.F., Ali Akman, M., Bahceci, M., 2023. Conception rates following assisted reproduction in poor responder patients: a retrospective study in 300 consecutive cycles. Reprod Biomed Online 6, 439–443. https://doi.org/10.1016/S1472-6483(10)62164-5.
Van Rooij, I.A.J., Broekmans, F.J.M., Te Velde, E.R., Fauser, B.C.J.M., Bancsi, L.F.J.M.M., De Jong, F.H., Themmen, A.P.N., 2002. Serum anti-Müllerian hormone levels: a novel measure of ovarian reserve. Hum Reprod 17, 3065–3071. https://doi.org/10.1093/HUMREP/17.12.3065.
Wang, B., Liu, W., Liu, Y., Zhang, W., Ren, C., Guan, Y., 2021. What Does Unexpected Suboptimal Response During Ovarian Stimulation Suggest, an Overlooked Group? Front Endocrinol (Lausanne) 12. https://doi.org/10.3389/FENDO.2021.795254.
Yong, P.Y.K., Brett, S., Baird, D.T., Thong, K.J., 2003. A prospective randomized clinical trial comparing 150 IU and 225 IU of recombinant follicle-stimulating hormone (Gonal-F*) in a fixed-dose regimen for controlled ovarian stimulation in in vitro fertilization treatment. Fertil Steril 79, 308–315. https://doi.org/10.1016/S0015-0282(02)04583-1.

Received 23 September 2023; received in revised form 16 January 2024; accepted 23 January 2024.

RBMO

ELSEVIER

ARTICLE

Cytosine–guanine–guanine repeats of FMR1 gene negatively affect ovarian reserve and response in Chinese women

BIOGRAPHY

Minyue Dong, MD, PhD, is Chief Professor of the Department of Reproductive Genetics, Women's Hospital, School of Medicine, Zhejiang University. His clinical practice focuses on genetic counselling, prenatal diagnosis and preimplantation genetic diagnosis. His research interests include genes responsible for genetic diseases and pathogenesis of diseases during pregnancy.

Xinyang Jin[a], Wenshan Zeng[a], Yanfei Xu[a], Pengzhen Jin[a,b], Minyue Dong[a,b,*]

KEY MESSAGE

A negative relationship was found between cytosine–guanine–guanine (CGG) repeats of *FMR1* gene and serum anti-Müllerian hormone, oestradiol, antral follicle count and oocyte yield in Chinese women. Mediation analysis indicated that ovarian function mediated the statistical association between CGG repeats and ovarian response.

ABSTRACT

Research question: Do cytosine–guanine–guanine (CGG) repeats of the *FMR1* gene affect ovarian function, ovarian response and assisted reproductive technology (ART) outcomes in Chinese women?

Design: A retrospective cohort study of 5869 women who underwent 8932 ART cycles at Women's Hospital, School of Medicine, Zhejiang University between January 2018 and June 2021. Basic hormone level, oocyte yield, embryo quality and the rate of live birth were considered as main outcome measures to evaluate the effects of CGG repeats on ovarian function, ovarian response and ART outcomes.

Results: The CGG repeats were negatively related to serum anti-Müllerian hormone (AMH), oestradiol, antral follicle count (AFC) and oocyte yield. A significant association was found between serum AMH, oestradiol and AFC even after age was controlled for. No statistically significant association, however, was found between CGG repeats and embryo quality or live birth rate. Ovarian function mediated the association between CGG repeats and ovarian response.

Conclusion: Increased CGG repeats on the *FMR1* gene were associated with diminished ovarian function and poor ovarian response, and ovarian function played an intermediary role in the relationship between CGG repeats and ovarian response.

INTRODUCTION

Fragile X mental retardation 1 (*FMR1*) gene is located at Xq27.3, and the dynamic expansion of cytosine–guanine–guanine (CGG) repeats in the 5′ untranslated region of exon 1 link to a variety of disorders, including fragile X syndrome (FXS), fragile X-associated tremor/ataxia syndrome, fragile X-associated primary ovarian insufficiency (FXPOI) and other fragile X-associated phenotypes (*Fu et al., 1991*).

The full mutation, an expansion of over 200 CGG repeats in *FMR1*, is responsible

[a] Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China
[b] Key Laboratory of Reproductive Genetics, Ministry of Education (Zhejiang University), Hangzhou, China

*Corresponding author. E-mail address: dongmy@zju.edu.cn (M. Dong). https://doi.org/10.1016/j.rbmo.2023.103779 1472-6483/
Declaration The authors report no financial or commercial conflicts of interest.

KEYWORDS

FMR1 gene
CGG repeats
assisted reproductive technology
ovarian function
ovarian response

for FXS, the most common inherited form of mental retardation in males, caused by the reduction in fragile X mental retardation protein mainly expressed in the brain and testis (*Fu et al., 1991*; *Lozano et al., 2014*). Gender plays an important role in FXS. All males and almost one-half of females with full mutation are exposed to FXS. For females, however, physical and behavioural symptoms of FXS are milder compared with males (*Rousseau et al., 1994*).

Conversely, the premutation, with CGG repeats of 55–200, is not associated with FXS but prone to fragile X-associated tremor/ataxia syndrome and FXPOI. In females, specialists in reproductive medicine have focused more on premutation for its association with primary ovarian insufficiency (POI), including 2–6% of isolated POI and 14% of familial POI. About 20% of females with premutation will develop FXPOI, with the most common CGG repeats ranging from 80–100 (*Man et al., 2017*; *Sherman, 2000*).

An association has been found between premutation of *FMR1* and ovarian function; however, the effects of CGG repeats on ovarian response, embryonic quality and assisted reproductive technology (ART) prognosis need to be elucidated. Some observations revealed that ovarian response lowers as CGG repeats increases (*Tsafrir et al., 2010*; *Elizur et al., 2014*; *Avraham et al., 2017*), but different findings have been reported (*Bibi et al., 2010*; *Ranisavljevic et al., 2020*). In addition, the sample sizes of the studies were relatively small.

Different ovarian functions and responses have been observed even in women with CGG repeats of normal or intermediate ranges. Poor ovarian function and low ovarian response were described in women of intermediate range (45–54 repeats), high normal range (35–44 repeats) as well as low normal range (<26 repeats) in studies with relatively small sample sizes (*Bretherick et al., 2005*; *Bodega et al., 2006*; *Gleicher et al., 2009*; *2013*; *Kushnir et al., 2014*; *Yang et al., 2016*; *Lu et al., 2017*; *Rehnitz et al., 2018*).

Cytosine–guanine–guanine repeats seem different among ethnicities. Compared with European and US populations, the frequency of premutation carrier was lower in Chinese population (*Seltzer et al., 2012*; *Hunter et al., 2014*; *Gao et al., 2020*) and the association between CGG repeats and POI less (*Guo et al., 2014*; *Man et al., 2017*; *Sherman, 2000*). When ethnicity is considered, studies of Chinese and East Asian populations were limited, and the findings differed among the relatively small samples investigated.

To elucidate the potential relationship between CGG repeats in *FMR1* gene and the ovarian function, ovarian response and the outcomes of ART in the Chinese population, research was conducted in a tertiary university hospital among a large sample population, and the relationship between CGG repeats and ovarian function, ovarian response, embryo quality and ART outcomes was analysed.

MATERIALS AND METHODS

Participants

A total of 13,206 women underwent screening of CGG repeats at Women's Hospital, School of Medicine, Zhejiang University between January 2018 and June 2021. The following women were excluded: women who did not complete ovarian function evaluation (n = 6828); women who had undergone ovarian surgery (n = 483); women of non-Chinese Han ethnicity (n = 23); and women who joined other clinical trials (n = 3). A total of 5869 women who had undergone 8932 ART cycles were recruited. Ovarian stimulation was used in 8932 ART cycles, fertilization was achieved in 8682 ART cycles and at least one embryo was cultured in 7179 ART cycles (FIGURE 1).

FMR1 grouping

The *FMR1* genotype was determined via triplet repeat polymerase chain reaction for size discrimination with capillary electrophoresis as previously described (*Monaghan et al., 2013*). The allele with greater CGG repeats was labeled as allele 2, and the opposite was allele 1 (*Gleicher et al., 2009*). All particpants were categorized by different amplification probabilities of CGG: low normal (<26 repeats), mid normal (26–34 repeats), high normal (35–44 repeats), intermediate (45–54 repeats) and premutation (55–200 repeats) (*Gleicher et al., 2014*; *Spector et al., 2021*).

Ovarian stimulation

Ovarian stimulation was carried out according to standard conventional protocols and physician recommendation. A total of 8932 ovarian stimulation cycles were recruited, and 2559 cycles of gonadotrophin releasing hormone agonist protocol (28.65%), 3862 cycles of antagonist protocol (43.24%), 685 cycles of minimal protocol (7.67%) and 1826 cycles of other protocols (20.44%) were carried out. The variations in the components of ovulation promotion protocols among *FMR1* groups are presented in FIGURE 2. The protocols have been described previously (*The ESHRE Guideline Group on Ovarian Stimulation et al., 2020*).

On days 2–5 of menstruation, basic FSH, LH, oestradiol and anti-Müllerian hormone (AMH) were assessed. Antral follicle count (AFC) was determined in both the ovaries by transvaginal ultrasonography. Diminished ovarian function was considered if any of the following was present: AFC lower than 5–7; FSH above 10 IU/l; FSH–LH ratio greater than 3.0–3.5; AMH less than 1.1 ng/ml (*Cohen et al., 2015*).

Ovarian response

Oocyte yield was the primary outcome of ovarian response. Total gonadotrophin dosage, gonadotrophin duration and peak oestradiol were also considered to be indices of ovarian response. Moreover, the follicular output rate (FORT) index, the ovarian sensitivity index (OSI) and follicle-to-oocyte index (FOI) were analysed. The FORT, OSI and FOI were calculated as follows: FORT = the number of follicles >14 mm on the day of trigger shot/AFC x 100%, OSI = (oocyte yield/total gonadotrophin dose) × 1000, FOI = oocyte yield/AFC x 100%. The methods for subgrouping these three indices were carried out as described previously (*Gallot et al., 2012*; *Alviggi et al., 2018*; *Revelli et al., 2020*). The low FOR, FOI and OSI suggested a hypo-response profile.

Embryo quality

The Gardner scoring system and the Istanbul consensus were used to evaluate embryo quality (*Schoolcraft et al., 1999*; *Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology et al., 2011*).

Assisted reproductive technology outcomes

Live birth rate was used as the primary outcome, and rates of biochemical pregnancy, clinical pregnancy and miscarriage as secondary outcomes.

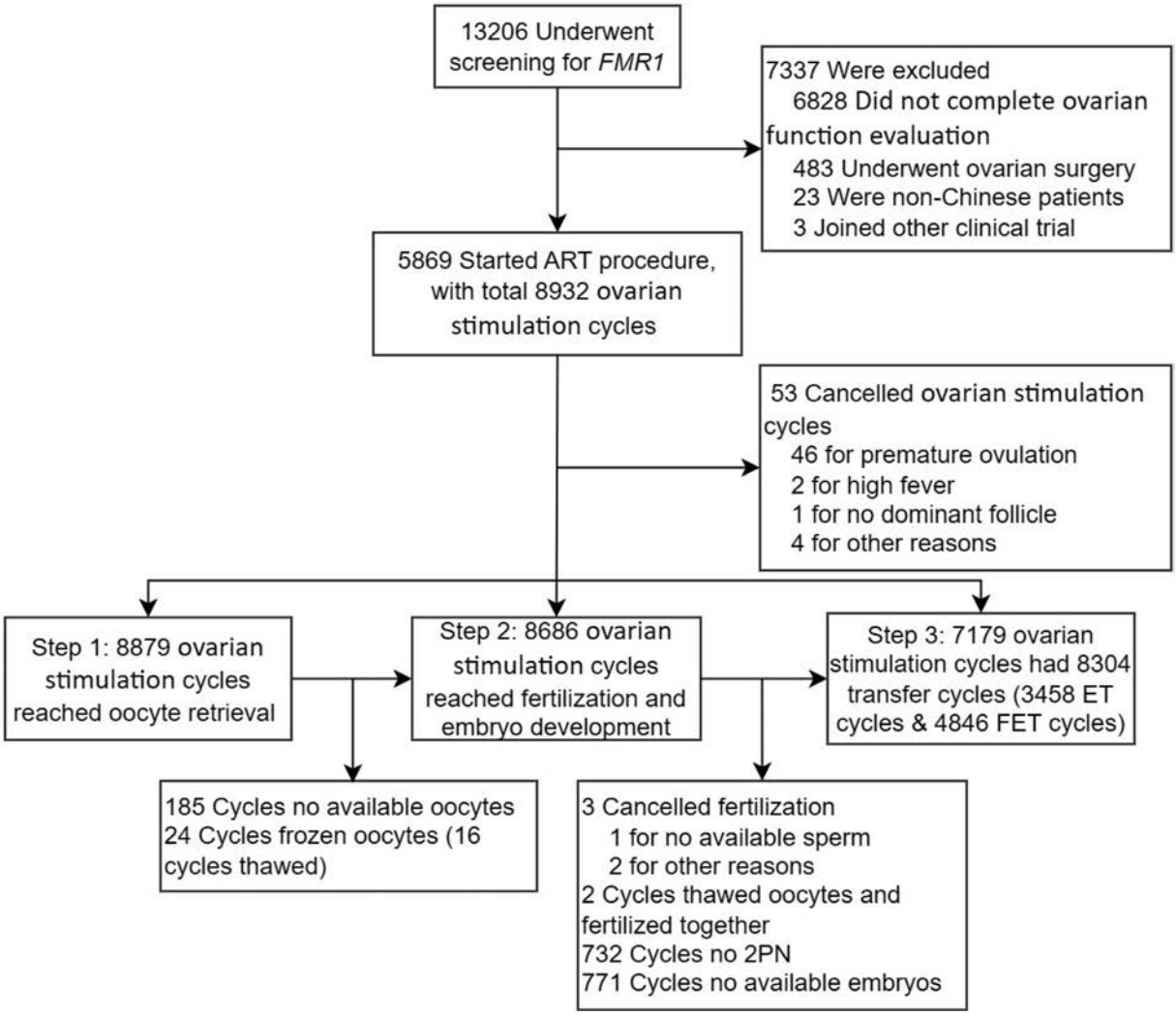


FIGURE 1 Study flowchart. Assisted reproductive technology (ART), normal fertilization (2PN), fresh embryo transfer (ET), frozen embryo transfer (FET).

Statistical analysis

The Statistical Package for the Social Sciences version 21.0 (SPSS Inc., Chicago, IL, USA) and the R Project for Statistical Modeling version 4.3.0 (R Core Team; Vienna, Austria) were used for statistical analysis. The Kolmogorov–Smirnov test was used to examine data normality. The skewed data were presented as medians and quartiles. Kruskal–Wallis, chi-squared and Fisher's exact tests were used where appropriate. Generalized estimating equations (GEE) were used to evaluate the association between CGG repeats and outcomes, and the results were presented as the mean difference or odds ratios with 95% confidence interval. Moreover, mediation analysis was conducted to explore the relationship between CGG repeats, ovarian function and ovarian response according to methods developed by *Iacobucci (2012)*. $P < 0.05$ was considered statistically significant. The hypothesized mediator model is presented in FIGURE 3.

Ethical approval

The Institutional Review Board of Women's Hospital, School of Medicine, Zhejiang University approved the collection of information from the medical records (IRB-20220066-R, 10 December 2021).

RESULTS

The characteristics are presented in TABLE 1. No significant differences were found in age ($P = 0.261$) and body mass index ($P = 0.118$) among the groups, although premutation group had a non-signicant trend towards higher age and body mass index than other groups. The premutation group had more secondary infertility than other groups but not significantly ($P = 0.142$).

Basic hormone levels depending on the CGG repeats are presented in TABLE 2. Oestradiol levels ($P = 0.074$) and FSH ($P = 0.928$) were not significantly different among the groups; however, significant differences were observed in concentrations of AMH ($P = 0.001$) and LH ($P = 0.009$), numbers of AFC ($P = 0.009$) and ovarian function ($P = 0.005$). The AMH and LH concentrations, AFC and ovarian function were significantly decreased in premutation group compared with other groups. Diminished ovarian function was observed in 1500 women who received 3168 ovarian stimulation cycles.

As shown in TABLE 3, significant differences were observed in days of gonadotrophin

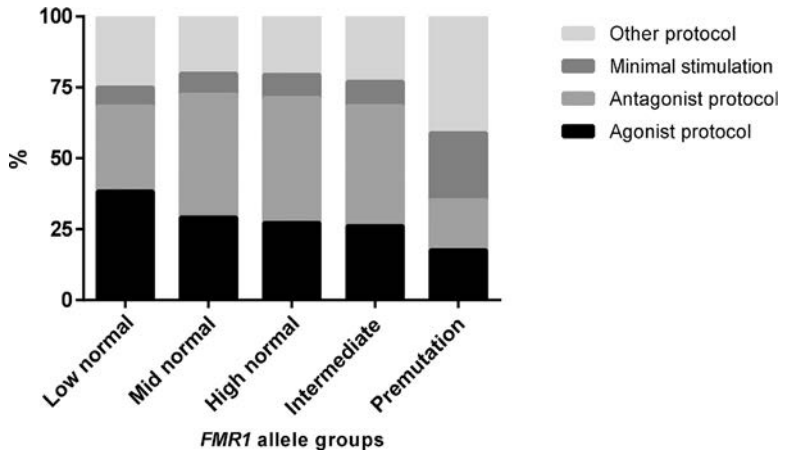


FIGURE 2 Incidence of ovarian stimulation protocols among the *FMR1* groups. CCG, cytosine–guanine– guanine.

use ($P = 0.009$), peak oestradiol levels ($P = 0.001$) and oocyte retrieval ($P = 0.001$) among the groups. The peak oestradiol levels and oocyte retrievals were significantly reduced in the premutation group compared with the other groups. No significant differences were observed in rates of normal fertility and good-quality embryos (TABLE 4). No significant differences were found in rates of chemical pregnancy, clinical pregnancy, miscarriage and live birth (TABLE 5).

The CGG repeats were significantly negatively related to AMH ($P = 0.006$), oestradiol ($P = 0.012$) and AFC ($P < 0.001$), indicating its association with diminished ovarian function. These associations remained after adjusting for age (TABLE 6). Furthermore, the FMR1 groupings for analysis of ovarian function in GEE model were also considered. Compared with the mid-normal group, AMH ($P < 0.001$) and oestradiol ($P = 0.015$) concentrations were significantly decreased in the premutation group regardless of the adjustment for age. At the same time, AFC was significantly decreased in the high-normal group ($P = 0.016$) and the premutation group ($P = 0.005$), and only decreased in the high-normal group ($P = 0.007$) when age was adjusted (TABLE 7).

The CGG repeats were significantly negatively associated with gonadotrophin duration ($P = 0.001$) and oocyte yield ($P = 0.013$) in the GEE model when age and ovarian function were unadjusted. Only the relationship between CGG repeats and gonadotrophin duration ($P = 0.007$), however, was observed after the adjustment for age and ovarian function (TABLE 8).

Compared with other ovarian response indicators, CGG repeats were not correlated with gonadotrophin dosage in the GEE model regardless of the adjustment of age and ovarian function. Furthermore, among the ovarian response indices, whatever the age, adjusted or not, no association was observed between CGG repeats and OSI, FORT, FOI (TABLE 8).

A significantly negative relationship was found between CGG repeats and fertilization rate only in the intracytoplasmic sperm injection (ICSI) group in the unadjusted GEE model ($P = 0.023$) and the age- and ovarian function-adjusted GEE model ($P = 0.016$). No relationship between CGG repeats was observed in the fertilization rate in the IVF and preimplantation genetic testing (PGT) groups, and the rate of good-quality embryo in IVF, ICSI and PGT groups (TABLE 9).

Because frozen embryo transfers (FET) were used more flexibly and involved multiple ovarian stimulation cycles, FET cycles were not included in the multifactorial analysis. Among the fresh embryo transfer cycles ($n = 3458$), no association was detected between CGG repeats and the rates of biochemical pregnancy, clinical pregnancy, miscarriage and live birth (TABLE 9).

To illustrate the mediating effect of ovarian function on the relationship between CGG repeats and ovarian response, a mediation analysis was conducted. Ovarian function accounted for 36.36% of the effect of CGG repeats on gonadotrophin duration whereas it accounted for nearly all of the effect on oocyte yield (TABLE 10).

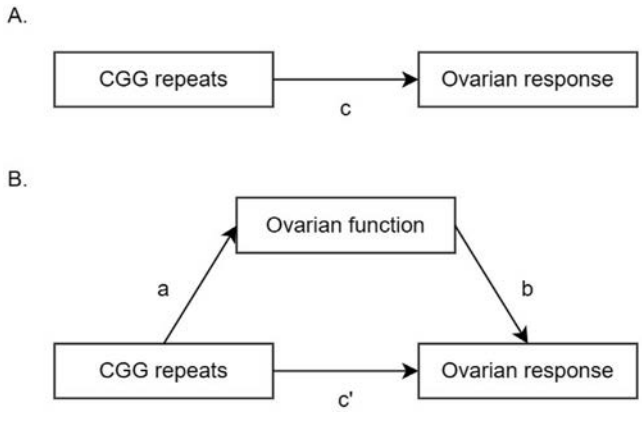


FIGURE 3 Path diagram of mediator model of ovarian function in cytosine–guanine– guanine (CGG) repeats and ovarian response. (A) The (c) path represents the total exposure (CGG repeats) -outcome (ovarian response, including oocyte yield and gonadotrophin duration) effect; (B) the (a) path represents the exposure-mediator effect, the (b) path represents the mediator (ovarian function) -outcome effect, and the (c′) path represents the direct exposure -outcome effect.

DISCUSSION

In the present study, the effects of CGG repeats of *FMR1* on ovarian function, ovarian response and ART outcomes were evaluated in a large cohort of Chinese patients. It has been shown that CGG repeats were associated with poor ovarian function and low ovarian response, but not embryo quality and ART outcomes. In addition, we explored that ovarian function acted as a mediator between CGG repeats and ovarian response.

TABLE 1 PARTICIPANT CHARACTERISTICS

Characteristics	*FMR1* allele groups					*P*-value
	Normal repeats			Intermediate (*n* = 83)	Premutation (*n* = 6)	
	Low normal	Mid normal	High normal			
	(*n* = 43)	(*n* = 4132)	(*n* = 1605)			
Age, years	31.00 (28.00–34.00)	31.00 (28.00–34.00)	30.00 (28.00–34.00)	31.00 (28.00–33.00)	35.00 (31.25–39.50)	0.261
Woman BMI, kg/m^2	21.48 (19.14–22.77)	21.48(19.72–23.53)	21.48 (19.68–23.58)	21.23 (19.96-23.62)	24.64 (22.34–25.18)	0.118
Infertility type, *n* (%)						0.851
Primary	23 (53.49)	2065 (49.98)	790 (49.22)	39 (46.99)	2 (33.33)	
Secondary	20 (46.51)	2067 (50.02)	815 (50.78)	44 (53.01)	4 (66.67)	

Non-normally distributed data are reported as median values (with first and third quartile in parenthesis). Categorical variables are expressed as number (percentage). Continuous variables are calculated by Kruskal–Wallis H test; categorical variables are calculated by Fisher's exact chi-squared test. BMI, body mass index.

TABLE 2 ASSOCIATIONS BETWEEN *FMR1* ALLELE GROUPS AND OVARIAN FUNCTION MARKERS

Ovarian function markers	*FMR1* allele groups					*P*-value
	Normal repeats			Intermediate (*n* = 130)	Premutation (*n* = 17)	
	Low normal	Mid normal	High normal			
	(*n* = 60)	(*n* = 6305)[a]	(*n* = 2415)			
Basic hormone level						
FSH, IU/l	7.18 (5.69–9.17)	6.90 (5.73–8.55)	6.88 (5.78–8.50)	6.92 (5.77–7.99)	6.39 (6.29–7.83)	0.928
LH, IU/l	5.43 (3.57–6.67)[b]	4.90 (3.61–6.58)[b]	4.92 (3.57–6.60)[b]	4.90 (3.58–6.60)[b]	2.44 (2.23–4.42)	0.009
Oestradiol, pmol/l	108.40 (78.20–135.80)	115.35 (83.27–155.43)	113.10 (80.92–151.70)	116.50 (82.42–149.53)	75.64 (60.71–128.80)	0.074
AMH, ng/ml	2.65 (1.14–4.56)[b]	2.28 (0.99–4.20)[b]	2.15 (0.97–4.10)[b]	2.50 (1.20–4.39)[b]	0.65 (0.65–1.28)	0.001
AFC, *n*	11.50 (7.13–17.00)[b]	11.00 (6.50–15.00)[b]	11.00 (6.00–15.00)	11.00 (6.00–16.00)	5.00 (4.00–10.00)	0.009
Ovarian function, *n* (%)						0.005
Normal	43 (71.67)[b]	4086 (64.81)[b]	1547 (64.06)[b]	79 (60.77)[b]	4 (23.53)	
Diminished	17 (28.33)	2219 (35.19)	868 (35.94)	51 (39.23)	13 (76.47)	

Non-normally distributed data are reported as median values (with first and third quartile in parenthesis); categorical variables are expressed as number (percentage). Continuous variables are calculated by Kruskal–Wallis H test; categorical variables are calculated by chi-squared test.

[a] The basic FSH, LH, oestradiol, anti-Müllerian hormone (AMH) and antral follicle count (AFC) of five cases were not recorded in the system.

[b] $P < 0.05$ compared with premutation group.

n number of ovarian stimulation cycles.

First, we found that, as CGG repeats increased, the risk of low AMH, oestradiol, AFC and ovarian dysfunction increased. Oestradiol and AMH were significantly lower in premutation but not in the intermediate range, low-normal range or high-normal range compared with the mid-normal range. Nevertheless, in the GEE model, we failed to prove the premutation was directly associated with ovarian dysfunction, possibly for the limited sample size in premutation. As with our findings, *Tsafrir et al. (2010)* reported that premutation carriers had decreased AFC and AMH compared with the normal range. *Morin et al. (2016)* and *Schufreider et al. (2015)*, however, proposed that CGG repeats were not associated with AMH and AFC, possibly because no premutation carriers were included in their observations. The exact mechanism of ovarian dysfunction related to the premutation is not fully understood. Some studies have partially explained the effect of CGG repeats in *FMR1* gene on ovarian function, including, but not limited to, excessive mRNA accumulation, reduced expression of Fragile X mental retardation protein, the presence of FMRpolyG in ovarian inclusions and altered histone modifications in the *FMR1* gene (*Tassone et al., 2000*; *Buijsen et al., 2016*; *Wang et al., 2018*).

We also demonstrated a weak tendency of a decrease in oocyte yield and a significant decrease in gonadotrophin duration with the increase in CGG repeats, which was in accordance with previously published studies on ovarian response with CGG repeats (*Gleicher et al., 2009*; *2013*; *Banks et al., 2016*). *Morin et al. (2016)* and *Cogendez et al. (2022)*, however, failed to demonstrate the relationship between CGG repeats and ovarian response, possibly because premutation carriers were included in the current investigation but not theirs.

Furthermore, we investigated the relationship between CGG repeats,

TABLE 3 ASSOCIATIONS BETWEEN FMR1 ALLELE GROUPS AND OVARIAN RESPONSE MARKERS

Ovarian response markers	*FMR1* allele groups					*P*-value
	Normal repeats			Intermediate (*n* = 130)	Premutation (*n* = 17)	
	Low normal	Mid normal	High normal			
	(*n* = 60)	(*n* = 6310)	(*n* = 2415)			
Gonadotrophin duration, days	10.00 (8.00–11.00)[a]	10.00 (8.00–11.00)[a]	9.00(8.00–11.00)[a]	9.00 (8.00–11.00)[a]	7.00 (2.50–9.00)	0.009
Gonadotrophin dosage, IU	2025.00 (1500.00–2250.00)	2025.00 (1500.00–2437.50)	1987.50 (1500.00–2400.00)	1950.00 (1350.00–2475.00)	1400.00 (375.00–2025.00)	0.264
Peak oestradiol, pmol/l	8775.50 (4261.25–15705.25)[a]	7955.00 (4076.50–13470.00)[a]	7685.00 (4083.00–13539.50)[a]	7349.50 (3879.25–12175.50)[a]	1874.00 (1074.95–6455.00)	0.001
Oocytes retrieved, *n*	8.00 (3.00–14.00)[a]	8.00 (4.00–14.00)[a]	8.00 (4.00–13.00)[a]	8.00 (4.00–13.00)[a]	2.00 (1.00–6.00)	0.001
FORT group,[b] *n* (%)						
Low FORT	21 (37.50)	1890 (33.05)	741 (33.36)	39 (32.23)	8 (50.00)	0.642
Average FORT	13 (23.21)	1946 (34.03)	751 (33.81)	38 (31.40)	5 (31.25)	
High FORT	22 (39.29)	1882 (32.91)	729 (32.82)	44 (36.36)	3 (18.75)	
OSI group,[b] *n* (%)						0.224
Low OSI	14 (24.56)	1491 (24.36)	580 (24.70)	31(24.41)	8 (57.14)	
Intermediate OSI	31 (54.39)	3064 (50.07)	1187 (50.55)	66(51.97)	6 (42.86)	
High OSI	12 (21.05)	1565 (25.57)	581 (24.74)	30(23.62)	0 (0.00)	
FOI group,[b] *n* (%)						0.080
≤50%	15 (26.79)	1555 (27.06)	582 (25.84)	36(29.75)	9 (56.25)	
>50%	41 (73.21)	4191 (72.94)	1670 (74.16)	85(70.25)	7 (43.75)	

Non-normally distributed data are reported as median values (with first and third quartile in parenthesis); categorical variables are expressed as the number (percentage). R Project was used to assess differences in follicular output rate (FORT) and ovarian sensitivity index (OSI) among *FMR1* genotypes; categorical variables were calculated by Fisher's exact chi-squared test.

[a] $P < 0.05$ compared with premutation group.

[b] The absence of antral follicle count or gonadotrophin use led to the failure of follicle-to-oocyte index (FOI), ovarian sensitivity index (OSI) and FORT for some cycles and, subsequently, the discrepancy of sample size.

n, number of ovarian stimulation cycles.

TABLE 4 ASSOCIATIONS BETWEEN *FMR1* ALLELE GROUPS AND EMBRYO QUALITY MARKERS

Embryo quality markers		*FMR1* allele groups					*P*-value
		Normal repeats			Intermediate (*n* = 126)	Premutation (*n* = 14)	
		Low normal	Mid normal	High normal			
		(*n* = 56)	(*n* = 6129)	(*n* = 2357)			
Rate of normal fertilization, %							
	IVF	57.14 (32.29–84.52)	64.29 (44.83–83.33)	66.67 (42.86–83.33)	63.96 (42.61–86.43)	30.00 (0.00–92.50)	0.808
	ICSI	72.50 (44.44–100.00)	66.67 (42.86–91.83)	62.50 (40.00–88.89)	75.00 (36.67–100.00)	0.00 (0.00–100.00)	0.422
	PGT	NA	72.73 (55.90–87.50)	67.54 (60.00–84.89)	77.50 (64.69–87.29)	NA	0.698
Rate of good-quality embryo, %		25.00 (0.00-57.14)	33.33 (0.00–50.00)	33.33 (0.00–50.00)	25.00 (0.00–50.00)	66.67 (0.00–90.00)	0.690

No individuals in the low-normal group and premutation group underwent preimplantation genetic testing (PGT).

ICSI, intracytoplasmic sperm injection; *n*, number of ovarian stimulation cycles in which fertilization was completed.

TABLE 5 ASSOCIATIONS BETWEEN *FMR1* ALLELE GROUPS AND PREGNANCY OUTCOMES

Pregnancy outcomes	*FMR1* allele groups					*P*-value
	Normal repeats			Intermediate (*n* = 113)	Premutation (*n* = 5)	
	Low normal	Mid normal	High normal			
	(*n* = 57)	(*n* = 5909)	(*n* = 2220)			
Transfer type, *n* (%)						0.990
Fresh	22 (38.60)	2462 (41.67)	926 (41.71)	46 (40.71)	2 (40.00)	
Frozen	35 (61.40)	3447 (58.34)	1294 (58.29)	67 (59.29)	3 (60.00)	
Biochemical pregnancy rate, *n* (%)	26 (45.61)	3192 (54.02)	1237 (55.72)	65 (57.52)	2 (40.00)	0.330
Clinical pregnancy rate, *n* (%)	26 (45.61)	2960 (50.09)	1156 (52.07)	60 (53.10)	2 (40.00)	0.449
Miscarriage rate, *n* (%)	3 (11.54)	495 (16.72)	180 (15.57)	10 (16.67)	0 (0.00)	0.856
live birth rate, *n* (%)	23 (40.35)	2456 (41.56)	974 (43.87)	50 (44.25)	2 (40.00)	0.421

Categorical variables were calculated by Fisher's exact chi-squared test.

n, number of transfer cycles.

TABLE 6 ASSOCIATIONS BETWEEN CYTOSINE–GUANINE–GUANINE REPEATS AND OVARIAN FUNCTION MARKERS IN GENERALIZED ESTIMATING EQUATION MODELS

Ovarian function markers	CGG repeats			
	Unadjusted GEE model Mean difference (95% CI)	*P*-value	Adjusted GEE model Mean difference (95% CI)	*P*-value
FSH	0.008 (−0.014 to 0.029)	0.483	0.004 (−0.018 to 0.027)	0.695
LH	−0.014 (−0.029 to 0.001)	0.060	−0.012 (−0.026 to 0.001)	0.072
Oestradiol	−0.417 (−0.742 to −0.093)	0.012	−0.438 (−0.751 to −0.125)	0.006
AMH	−0.019 (−0.033 to −0.005)	0.006	−0.014 (−0.027 to −0.001)	0.030
AFC	−0.052 (−0.079 to −0.025)	<0.001	−0.041 (−0.069 to −0.014)	0.003
	Odds ratio (95% CI)		Odds ratio (95% CI)	
Ovarian function	1.017 (1.005 to 1.028)	0.004	1.013 (1.001 to 1.025)	0.039

Adjusted model accounted for patient age.

AFC, antral follicle count; AMH, anti-Mullerian hormone; CGG, cytosine–guanine–guanine; GEE, generalized estimating equation.

TABLE 7 ASSOCIATIONS BETWEEN *FMR1* ALLELE GROUPS AND OVARIAN FUNCTION MARKERS IN GENERALIZED ESTIMATING EQUATION MODELS

Ovarian function markers		*FMR1* alleles groups							
		Low normal		High normal		Intermediate		Premutation	
	Corrections	Mean difference (95% CI)	*P*-value	Mean difference (95% CI)	*P*-value	Mean difference (95% CI)	*P*-value	Mean difference (95% CI)	*P*-value
FSH	Unadjusted	0.674 (−0.581 to 1.928)	0.293	0.129 (−0.126 to 0.385)	0.321	−0.326 (−0.851 to 0.199)	0.224	2.519 (−3.244 to 8.281)	0.392
	Adjusted	0.681 (−0.586 to 1.948)	0.292	0.138 (−0.111 to 0.387)	0.276	−0.355 (−0.847 to 0.136)	0.156	1.754 (-4.092,7.601)	0.556
LH	Unadjusted	0.665 (−0.427 to 1.757)	0.233	−0.013 (−0.195 to 0.169)	0.889	0.134 (−0.618 to 0.887)	0.727	−1.389 (−3.716 to 0.938)	0.242
	Adjusted	0.644 (−0.432 to 1.720)	0.241	−0.017 (−0.198 to 0.165)	0.856	0.140 (−0.602 to 0.881)	0.712	−1.041 (−3.165 to 1.082)	0.337
Oestradiol	Unadjusted	−10.526 (−30.508 to 9.456)	0.302	−6.593 (−12.606 to −0.579)	0.032	−0.669 (−15.529 to 14.192)	0.930	−28.027 (−50.661 to −5.394)	0.015
	Adjusted	−10.284 (−30.299 to 9.662)	0.312	−6.486 (−12.441 to −0.531)	0.033	−0.750 (−15.554 to 14.055)	0.921	−32.388(-54.508,-10.268)	0.004
AMH	Unadjusted	0.277 (−0.665 to 1.128)	0.565	−0.119 (−0.293 to 0.055)	0.179	0.079 (−0.545 to 0.703)	0.804	−2.091 (−2.789 to −1.393)	<0.001
	Adjusted	0.255 (−0.649 to 1.160)	0.580	−0.109 (−0.280 to 0.062)	0.211	0.138 (−0.457 to 0.732)	0.650	−1.130 (−1.602 to −0.657)	<0.001
AFC	Unadjusted	0.045 (−1.943 to 2.033)	0.965	−0.425 (−0.769 to −0.081)	0.016	0.020 (−1.413 to 1.453)	0.978	−4.211 (−7.167 to −1.255)	0.005
	Adjusted	0.053 (−1.901 to 2.007)	0.958	−0.439 (−0.758 to −0.120)	0.007	0.124 (−1.195 to 1.444)	0.854	−2.010 (−4.311 to 0.290)	0.087
		Odds ratio (95% CI)		Odds ratio (95% CI)		Odds ratio (95% CI)		Odds ratio (95% CI)	
Ovarian function	Unadjusted	0.820 (0.403 to 1.670)	0.585	1.107 (0.976 to 1.257)	0.115	1.266 (0.794 to 2.020)	0.322	2.956 (0.595 to 14.695)	0.185
	Adjusted	0.781 (0.360 to 1.691)	0.530	1.104 (0.964 to1.265)	0.153	0.991 (0.576 to 1.704)	0.974	1.805 (0.471 to 6.921)	0.389

Adjusted model accounted for patient age; mid-normal group used as the reference.

AFC, antral follicle count; AMH, anti-Müllerian hormone.

TABLE 8 ASSOCIATIONS BETWEEN CYTOSINE–GUANINE–GUANINE REPEATS AND OVARIAN RESPONSE MARKERS IN GENERALIZED ESTIMATING EQUATION MODELS

Ovarian response markers	CGG repeats			
	Unadjusted GEE model Mean difference (95% CI)	*P*-value	Adjusted GEE model Mean difference (95% CI)	*P*-value
Gonadotrophin duration	−0.027 (−0.044 to −0.010)	0.001	−0.021 (−0.035 to -0.006)	0.007
Gonadotrophin dosage	−3.916 (−8.613 to 0.781)	0.102	−3.365 (−8.086 to 1.356)	0.162
Oocytes retrieved	−0.043 (−0.077 to −0.009)	0.013	−0.011 (−0.033 to 0.011)	0.340
FOI	−0.175 (−0.465 to 0.116)	0.239	−0.133 (−0.394 to 0.128)	0.318
OSI	0.019 (−0.075 to 0.112)	0.697	0.031 (−0.061 to 0.123)	0.510
FORT	−0.014(−0.334 to 0.307)	0.934	−0.010 (−0.327 to 0.307)	0.951

Gonadotrophin dosage and oocyte yield adjusted for age and ovarian function.

Follicle-to-oocyte index (FOI), ovarian sensitivity index (OSI) and follicular output rate index (FORT) variables adjusted for age.

TABLE 9 ASSOCIATIONS BETWEEN CYTOSINE–GUANINE–GUANINE REPEATS WITH EMBRYO QUALITY AND PREGNANCY OUTCOME IN GENERALIZED ESTIMATING EQUATION MODELS FOR FRESH EMBRYOS ONLY

Embryo quality markers/ pregnancy outcome	CGG repeats			
	Unadjusted GEE model Mean difference (95% CI)	*P*-value	Adjusted GEE model Mean difference (95% CI)	*P*-value
Rate of normal fertilization				
IVF	−0.001 (−0.003 to 0.001)	0.428	−0.001 (−0.004 to 0.001)	0.412
ICSI	−0.001 (−0.002 to 0.000)	0.023	−0.001 (−0.002 to 0.000)	0.016
PGT	0.007 (−0.004 to 0.019)	0.224	0.008 (−0.004 to 0.021)	0.188
Rate of good quality embryo	0.000 (−0.001 to 0.002)	0.653	0.000 (−0.002 to 0.002)	0.844
	Odds ratio (95% CI)		Odds ratio (95% CI)	
Biochemical pregnancy	0.997 (0.982 to 1.012)	0.695	0.998 (0.983 to 1.013)	0.747
Clinical pregnancy	0.997 (0.982 to 1.012)	0.704	0.998 (0.983 to 1.013)	0.764
Miscarriage	0.989 (0.961 to 1.018)	0.440	0.988 (0.960 to 1.018)	0.433
Live birth	1.001 (0.986 to 1.017)	0.880	1.003 (0.987 to 1.018)	0.747

Adjusted model of cytosine–guanine–guanine (CGG) repeats with embryo quality accounted for patient age and ovarian function; adjusted model of CGG repeats with assisted reproduction outcomes accounted for patient age, ovarian function and number of embryos transferred.

ICSI, intracytoplasmic sperm injection; PGT, preimplantation genetic diagnosis.

ovarian function and ovarian response. We observed that CGG repeats affected ovarian response by influencing ovarian function. In other words, CGG repeats were related to ovarian response, and this relationship was achieved partially through ovarian function. To the best of our knowledge, this is the first description of the mediating effect of ovarian function in CGG repeats and ovarian response.

Finally, no significant relationship was found between CGG repeats, embryo quality and pregnancy outcomes, in accordance with the findings of *Lledo et al. (2012)* and *Banks et al. (2016)*. Unexpectedly, we found a negative association between CGG repeats and fertilization rate in the ICSI group. No relationship with fertilization rate was observed in IVF and PGT. On the basis of these results, we believe that CGG repeats were not related to embryo quality. This might be because the indication for ICSI was primarily related to problems with spermatozoa. In other words, this observation was affected by the quality of the spermatozoa, and, therefore, a relationship between CGG repeats and fertilization rate, and the quality of the embryo, could not be established. It was initially thought that results were based on transfer cycles; however, CGG repeats had affected ovarian function and retrieved oocytes, leading to a special clinical decision to improve outcome. On the positive side, this also meant that clinically similar embryo quality and pregnancy outcomes could be achieved through adjustments to the ovulation protocol and gonadotrophin use, as well as the choice of ART.

The premutation carrier frequency ranged from one in 291 to one in 130 in the European and US populations (*Seltzer et al., 2012*; *Hunter et al., 2014*). The premutation carrier frequency in Chinese women who had premature ovarian failure was two in 379 (*Guo et al., 2014*), in Chinese women of childbearing age one in 634 (*Gao et al., 2020*) and in pregnant women from Hong Kong, China, one in 1325 (*Cheng et al., 2017*). It seemed that

TABLE 10 MEDIATING EFFECTS OF OVARIAN FUNCTION IN CYTOSINE–GUANINE–GUANINE REPEATS AND OVARIAN RESPONSE

		Model 1: oocyte yield		Model 2: gonadotropin duration	
		Coefficient	95% CI	Coefficient (95% CI)	95% CI
CGG repeats	Total effect, c	−0.043[a]	−0.077 to −0.009	−0.027[a]	−0.044 to −0.010
	Direct effect, c′	−0.018	−0.043 to 0.008	−0.022[a]	−0.038 to −0.006
	Mediating effect, $Z_{Mediation}$	−0.119[a]	−0.188 to −0.049	−0.026[a]	−0.043 to −0.010
	Mediated, %	100.00		36.36	

[a] $P < 0.05$.

Mediating effect was determined by examining the significance of the Z_a*Z_b($Z_{Mediation}$) in R project using medci {RMediation}. Z_a = a/SE(a),Z_b = b/SE(b). Where SE(a), the standard error of estimate a; SE(b), the standard error of estimate b.

the premutation carrier frequency was lower in Chinese women.

In the present study, the premutation frequencies were six in 5869 but three in 1500 in women with poor ovarian function. Similar to previous studies, it showed a higher frequency of premutation carrier in those with poor ovarian function (*Guo et al., 2014*) and a lower frequency compared with European and US populations. This carrier frequency, however, seemed to be much lower compared with some previously published reports (*Guo et al., 2014*; *Gao et al., 2020*). The discrepancy in premutation carrier frequency might be associated with the population in which the observation was conducted.

To the best of our knowledge, the present study has the largest sample size of normal range, intermediate and premutation from which to investigate the effects of CGG repeats on ovarian function, ovarian response and ART outcomes in a Chinese population. Because of the low frequency of premutation in Chinese population, the sample size in premutation in our study was limited. We observed that premutation carriers were more prone to secondary infertility, but not significantly. The limited sample size, however, might affect the accuracy of extrapolated results; therefore, further studies with larger sample sizes of premutation carriers are needed.

In conclusion, we demonstrated that CGG repeats were related to ovarian function and ovarian response, while simultaneously influencing ovarian response by affecting ovarian function. Therefore, for infertility, *FMR1* gene screening is a good way to reduce or avoid birth defects in offspring, and also a useful test that can be used as a biological indicator for early assessment of ovarian reserve function, helping clinicians to eliminate the effect on embryo quality and pregnancy outcome through adjustment of assisted reproduction strategies, and to help women make their family planning in a timely way.

DATA AVAILABILITY

Data will be made available on request.

ACKNOWLEDGEMENTS

We wish to thank Professor Yunxian Yu and his team, who made suggestions on statistical analysis.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant numbers 82171848), Zhejiang Provincial Natural Science Foundation of China (LY22H110004).

REFERENCES

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, Balaban, B., Brison, D., Calderon, G., Catt, J., Conaghan, J., Cowan, L., Ebner, T., Gardner, D., Hardarson, T., Lundin, K., Cristina Magli, M., Mortimer, D., Mortimer, S., Munne, S., Royere, D., Scott, L., Smitz, J., Thornhill, A., Van Blerkom, J., Van Den Abbeel, E., 2011. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Human Reproduction 26, 1270–1283.

Alviggi, C., Conforti, A., Esteves, S.C., Vallone, R., Venturella, R., Staiano, S., Castaldo, E., Andersen, C.Y., De Placido, G., 2018. Understanding Ovarian Hypo-Response to Exogenous Gonadotropin in Ovarian Stimulation and Its New Proposed Marker—The Follicle-To-Oocyte (FOI) Index. Front. Endocrinol. 9, 589.

Avraham, S., Almog, B., Reches, A., Zakar, L., Malcov, M., Sokolov, A., Alpern, S., Azem, F., 2017. The ovarian response in fragile X patients and premutation carriers undergoing IVF—PGD: reappraisal. Human Reproduction 32, 1508–1511.

Banks, N., Patounakis, G., Devine, K., DeCherney, A.H., Widra, E., Levens, E.D., Whitcomb, B.W., Hill, M.J., 2016. Is *FMR1* CGG repeat length a predictor of in vitro fertilization stimulation response or outcome? Fertility and Sterility 105, 1537–1546.e8.

Bibi, G., Malcov, M., Yuval, Y., Reches, A., Ben-Yosef, D., Almog, B., Amit, A., Azem, F., 2010. The effect of CGG repeat number on ovarian response among fragile X premutation carriers undergoing preimplantation genetic diagnosis. Fertility and Sterility 94, 869–874.

Bodega, B., Bione, S., Dalprà, L., Toniolo, D., Ornaghi, F., Vegetti, W., Ginelli, E., Marozzi, A., 2006. Influence of intermediate and uninterrupted *FMR1* CGG expansions in premature ovarian failure manifestation. Human Reproduction 21, 952–957.

Bretherick, K.L., Fluker, M.R., Robinson, W.P., 2005. *FMR1* repeat sizes in the gray zone and high end of the normal range are associated with premature ovarian failure. Hum Genet 117, 376–382.

Buijsen, R.A.M., Visser, J.A., Kramer, P., Severijnen, E.A.W.F.M., Gearing, M., Charlet-Berguerand, N., Sherman, S.L., Berman, R.F., Willemsen, R., Hukema, R.K., 2016. Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a

role in fragile X-associated primary ovarian insufficiency. Hum. Reprod. 31, 158–168.
Cheng, Y.K., Lin, C.S., Kwok, Y.K., Chan, Y., Lau, T., Leung, T., Choy, K., 2017. Identification of fragile X pre-mutation carriers in the Chinese obstetric population using a robust *FMR1* polymerase chain reaction assay: implications for screening and prenatal diagnosis. Hong Kong Med J 23, 110–116.
Cogendez, E., Ozkaya, E., Eser, A.Ç., Eken, M., Karaman, A., 2022. Can *FMR1* CGG repeat lengths predict the outcome in ICSI cycles? Ginekol Pol 93, 735–741.
Cohen, J., Chabbert-Buffet, N., Darai, E., 2015. Diminished ovarian reserve, premature ovarian failure, poor ovarian responder—a plea for universal definitions. J Assist Reprod Genet 32, 1709–1712.
Elizur, S.E., Lebovitz, O., Derech-Haim, S., Dratviman-Storobinsky, O., Feldman, B., Dor, J., Orvieto, R., Cohen, Y., 2014. Elevated Levels of *FMR1* mRNA in Granulosa Cells Are Associated with Low Ovarian Reserve in *FMR1* Premutation Carriers. PLoS ONE 9, e105121.
Fu, Y.-H., Kuhl, D.P.A., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkert, A.J.M.H., Holden, J.J.A., Fenwick, R.G., Warren, S.T., Oostra, B.A., Nelson, D.L., Caskey, C.T., 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. Cell 67, 1047–1058.
Gallot, V., Berwanger Da Silva, A.L., Genro, V., Grynberg, M., Frydman, N., Fanchin, R., 2012. Antral follicle responsiveness to follicle-stimulating hormone administration assessed by the Follicular Output RaTe (FORT) may predict in vitro fertilization-embryo transfer outcome. Human Reproduction 27, 1066–1072.
Gao, F., Huang, W., You, Y., Huang, J., Zhao, J., Xue, J., Kang, H., Zhu, Y., Hu, Z., Allen, E.G., Jin, P., Xia, K., Duan, R., 2020. Development of Chinese genetic reference panel for Fragile X Syndrome and its application to the screen of 10,000 Chinese pregnant women and women planning pregnancy. Molec Gen & Gen Med 8, e1236.
Gleicher, N., Kim, A., Kushnir, V., Weghofer, A., Shohat-Tal, A., Lazzaroni, E., Lee, H.-J., Barad, D.H., 2013. Clinical Relevance of Combined FSH and AMH Observations in Infertile Women. The Journal of Clinical Endocrinology & Metabolism 98, 2136–2145.
Gleicher, N., Kushnir, V.A., Weghofer, A., Barad, D.H., 2014. How the *FMR1* gene became relevant to female fertility and reproductive medicine. Front. Genet. 5.
Gleicher, N., Weghofer, A., Oktay, K., Barad, D.H., 2009. Can the *FMR1* (Fragile X) Gene Serve As Predictor of Response to Ovarian Stimulation? Reprod. Sci. 16, 462–467.
Guo, T., Qin, Y., Jiao, X., Li, G., Simpson, J.L., Chen, Z.-J., 2014. *FMR1* Premutation Is an Uncommon Explanation for Premature Ovarian Failure in Han Chinese. PLoS ONE 9, e103316.
Hunter, J., Rivero-Arias, O., Angelov, A., Kim, E., Fotheringham, I., Leal, J., 2014. Epidemiology of fragile X syndrome: A systematic review and meta-analysis. American J of Med Genetics Pt A 164, 1648–1658.
Iacobucci, D., 2012. Mediation analysis and categorical variables: The final frontier. J Consum Psychol 22, 582–594.
Kushnir, V.A., Yu, Y., Barad, D.H., Weghofer, A., Himaya, E., Lee, H.-J., Wu, Y.-G., Shohat-Tal, A., Lazzaroni-Tealdi, E., Gleicher, N., 2014. Utilizing *FMR1* Gene Mutations as Predictors of Treatment Success in Human In Vitro Fertilization. PLoS ONE 9, e102274.
Lledo, B., Guerrero, J., Ortiz, J.A., Morales, R., Ten, J., Llacer, J., Gimenez, J., Bernabeu, R., 2012. Intermediate and normal sized CGG repeat on the *FMR1* gene does not negatively affect donor ovarian response. Human Reproduction 27, 609–614.
Lozano, R., Rosero, C.A., Hagerman, R.J., 2014. Fragile X spectrum disorders. IRDR 3, 134–146.
Lu, C.-L., Li, R., Chen, X.-N., Xu, Y.-Y., Yan, L.-Y., Yan, J., Zhang, Y.-Y., Jin, H.-Y., Zhang, W.-X., Qiao, J., Zhen, X.-M., 2017. The 'normal' range of *FMR1* triple CGG repeats may be associated with primary ovarian insufficiency in China. Reproductive BioMedicine Online 34, 175–180.
Man, L., Lekovich, J., Rosenwaks, Z., Gerhardt, J., 2017. Fragile X-Associated Diminished Ovarian Reserve and Primary Ovarian Insufficiency from Molecular Mechanisms to Clinical Manifestations. Front. Mol. Neurosci. 10, 290.
Monaghan, K.G., Lyon, E., Spector, E.B., 2013. ACMG Standards and Guidelines for fragile X testing: a revision to the disease-specific supplements to the Standards and Guidelines for Clinical Genetics Laboratories of the American College of Medical Genetics and Genomics. Genetics in Medicine 15, 575–586.
Morin, S.J., Tiegs, A.W., Franasiak, J.M., Juneau, C.R., Hong, K.H., Werner, M.D., Zhan, Y., Landis, J., Scott, R.T., 2016. *FMR1* gene CGG repeat variation within the normal range is not predictive of ovarian response in IVF cycles. Reproductive BioMedicine Online 32, 496–502.
Ranisavljevic, N., Hess, M., Castelli, C., Willems, M., Ferrieres-Hoa, A., Girardet, A., Anahory, T., 2020. Are ovarian response and pregnancy rates similar in selected *FMR1* premutated and mutated patients undergoing preimplantation genetic testing? J Assist Reprod Genet 37, 1675–1683.
Rehnitz, J., Alcoba, D.D., Brum, I.S., Dietrich, J.E., Youness, B., Hinderhofer, K., Messmer, B., Freis, A., Strowitzki, T., Germeyer, A., 2018. *FMR1* expression in human granulosa cells increases with exon 1 CGG repeat length depending on ovarian reserve. Reprod Biol Endocrinol 16, 65.
Revelli, A., Gennarelli, G., Biasoni, V., Chiadò, A., Carosso, A., Evangelista, F., Paschero, C., Filippini, C., Benedetto, C., 2020. The Ovarian Sensitivity Index (OSI) Significantly Correlates with Ovarian Reserve Biomarkers, Is More Predictive of Clinical Pregnancy than the Total Number of Oocytes, and Is Consistent in Consecutive IVF Cycles. JCM 9, 1914.
Rousseau, F., Heitz, D., Tarleton, J., MacPherson, J., Dahl, N., Barnicoat, A., Mathew, C., Maddalena, A., Spiegel, R., Schinzel, A., Marcos, J.A.G., Schorderet, D.F., Schaap, T., Maccioni, L., Russo, S., Jacobs, P.A., Schwartz, C., Mandel, J.L., 1994. A Multicenter Study on Genotype-Phenotype Correlations in the Fragile X Syndrome, Using Direct Diagnosis with Probe StB 12.3: The First 2,253 Cases. Am. J. Hum. Genet..
Schoolcraft, W.B., Gardner, D.K., Lane, M., Schlenker, T., Hamilton, F., Meldrum, D.R., 1999. Blastocyst culture and transfer: analysis of results and parameters affecting outcome in two in vitro fertilization programs. Fertility and Sterility 72, 604–609.
Schufreider, A., McQueen, D.B., Lee, S.M., Allon, R., Uhler, M.L., Davie, J., Feinberg, E.C., 2015. Diminished ovarian reserve is not observed in infertility patients with high normal CGG repeats on the fragile X mental retardation 1 (*FMR1*) gene. Hum. Reprod. 30, 2686–2692.
Seltzer, M.M., Baker, M.W., Hong, J., Maenner, M., Greenberg, J., Mandel, D., 2012. Prevalence of CGG expansions of the *FMR1* gene in a US population-based sample. Am. J. Med. Genet. 159B, 589–597.
Sherman, S.L., 2000. Premature ovarian failure in the fragile X syndrome. Am. J. Med. Genet. 97, 189–194.
Spector, E., Behlmann, A., Kronquist, K., Rose, N.C., Lyon, E., Reddi, H.V., 2021. Laboratory testing for fragile X, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). Genetics in Medicine 23, 799–812.
Tassone, F., Hagerman, R.J., Taylor, A.K., Gane, L.W., Godfrey, T.E., Hagerman, P.J., 2000. Elevated Levels of *FMR1* mRNA in Carrier Males: A New Mechanism of Involvement in the Fragile-X Syndrome. The American Journal of Human Genetics 66, 6–15.
The ESHRE Guideline Group on Ovarian Stimulation, Bosch, E., Broer, S., Griesinger, G., Grynberg, M., Humaidan, P., Kolibianakis, E., Kunicki, M., La Marca, A., Lainas, G., Le Clef, N., Massin, N., Mastenbroek, S., Polyzos, N., Sunkara, S.K., Timeva, T., Töyli, M., Urbancsek, J., Vermeulen, N., Broekmans, F., 2020. ESHRE guideline: ovarian stimulation for IVF/ICSI†. Human Reproduction Open 2020, hoaa009.
Tsafrir, A., Altarescu, G., Margalioth, E., Brooks, B., Renbaum, P., Levy-Lahad, E., Rabinowitz, R., Varshaver, I., Eldar-Geva, T., 2010. PGD for fragile X syndrome: ovarian function is the main determinant of success. Human Reproduction 25, 2629–2636.
Wang, Q., Barad, D.H., Darmon, S.K., Kushnir, V.A., Wu, Y.-G., Lazzaroni-Tealdi, E., Zhang, L., Albertini, D.F., Gleicher, N., 2018. Reduced RNA expression of the *FMR1* gene in women with low (CGGn<26) repeats. PLoS ONE 13, e0209309.
Yang, W., Fan, C., Chen, L., Cui, Z., Bai, Y., Lan, F., 2016. Pathological Effects of the *FMR1* CGG-Repeat Polymorphism (5-55 Repeat Numbers): Systematic Review and Meta-Analysis. Tohoku J. Exp. Med. 239, 57–66.

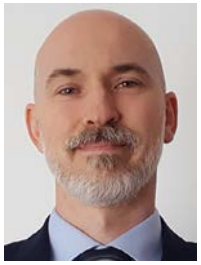
Received 13 October 2023; received in revised form 13 December 2023; accepted 20 December 2023.

RBMO

ELSEVIER

ARTICLE

Developing and validating a prediction model of live birth following single vitrified–warmed blastocyst transfer



BIOGRAPHY

Dr Piotr Sokol, MD, graduated in 2006 from Akademia Medyczna, Warsaw, and in 2012 completed specialty training in obstetrics and gynaecology at Hospital Clínic, Barcelona. Since then, he's been dedicated to assisted reproduction and is now a fertility consultant at Reproductive Medicine Department, Dexeus, Barcelona, with a special interest in male factor infertility.

Piotr Sokol[a,*,1], Elisabet Clua[a,1], María Carme Pons[a], Sandra García[a], Annalisa Racca[a], Thomas Freour[a,c,d], Nikolaos P Polyzos[a,b]

KEY MESSAGE

Blastulation day and blastocyst overall quality is linked to live birth rate (LBR) in single vitrified–thawed blastocyst transfer (SVBT) cycles. The prognostic model developed provides clinicians, patients and embryologists with precise LBR predictions for vitrified blastocysts, aiding in blastocyst selection and enhancing SVBT success.

ABSTRACT

Research question: Can the developed clinical prediction model offer an accurate estimate of the likelihood of live birth, involving blastocyst morphology and vitrification day after single vitrified–warmed blastocyst transfer (SVBT), and therefore assist clinicians and patients?

Study design: Retrospective cohort study conducted at a Spanish university-based reproductive medicine unit (2017–2021) including consecutive vitrified–warmed blastocysts from IVF cycles. A multivariable logistic regression incorporated key live birth predictors: vitrification day, embryo score, embryo ploidy status and clinically relevant variables, i.e. maternal age.

Results: The training set involved 1653 SVBT cycles carried out between 2017 and 2020; 592 SVBT cycles from 2021 constituted the external validation dataset. The model revealed that female age and embryo characteristics, including overall quality and blastulation day, is linked to live birth rate in SVBT cycles. Stratification by vitrification day and quality (from day-5A to day-6 C blastocysts) applied to genetically tested and untested embryos. The model's area under the curve was 0.66 (95% CI 0.64 to 0.69) during development and 0.65 (95% CI 0.61 to 0.70) in validation, denoting moderate discrimination. Calibration plots showed strong agreement between predicted and observed probabilities.

Conclusion: By incorporating essential predictors such as vitrification day, embryo morphology grade, age and preimplantation genetic testing for aneuploidy usage, this predictive model offers valuable guidance to clinicians and patients, enabling accurate forecasts of live birth rates for any given vitrified blastocyst within SVBT cycles. Additionally, it serves as a potentially indispensable laboratory tool, aiding in selecting the most promising blastocysts for optimal outcomes.

[a] Department of Obstetrics, Gynecology and Reproductive Medicine, Dexeus University Hospital, Barcelona, Spain
[b] Faculty of Health, University of Ghent, Ghent, Belgium
[c] Nantes Université, CHU Nantes, Inserm, CR2TI, F-44000 Nantes, France.
[d] CHU Nantes, Service de Medecine et Biologie de la Reproduction, F-44000 Nantes, France.
[1] Piotr Sokol and Elisabet Clua are co-authors.

*Corresponding author. E-mail address: piosok@dexeus.com (P. Sokol). https://doi.org/10.1016/j.rbmo.2024.103890 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEYWORDS
Live birth rate
prediction model
vitrification day
embryo grade
blastocyst

INTRODUCTION

Human assisted reproduction technology (ART) has evolved continuously since its introduction. It is marked by a few milestones in its relatively young history. Among them, the development of prolonged embryo culture up to the blastocyst stage and vitrification significantly changed practice and improved outcomes, facilitating the implementation of segmented cycles and single vitrified blastocyst transfer (SVBT). The International Committee Monitoring Assisted Reproductive Technologies and the European IVF Monitoring Consortium (EIM) recently reported sustained increase in the number of freeze-all cycles, frozen embryo transfer cycles (*Adamson et al., 2022*) and single embryo transfer (*European IVF Monitoring Consortium for the European Society of Human Reproduction and Embryology [ESHRE] et al., 2023*).

Selecting the embryo with the highest implantation potential for transfer (either fresh or frozen thawed) within a cohort is one of the most critical steps in IVF and intracytoplasmic sperm injection programmes to shorten the time to pregnancy and achieve a single and healthy live birth. Embryo quality is mainly determined by two important factors: morphology and ploidy status. In relation to morphology, even though some blastocyst scoring systems have been widely described (*Gardner and Schoolcraft, 1999*; *Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology et al., 2011*; *Cuevas Saiz et al., 2018*), their value as a predictor of live birth is still questionable, both for pre-freezing parameters (*Ahlstrom et al., 2013*; *Irani et al., 2017*; *Nazem et al., 2019*) and post-warming blastocyst characteristics (*Ahlstrom et al., 2013*; *Coello et al., 2017*; *Cimadomo et al., 2018*; *Giunco et al., 2021*; *Allen et al., 2022*). In some studies, blastocyst morphology predicted implantation (*Irani et al., 2019*; *Nazem et al., 2019*) but not in the study by Capalbo (2014). Furthermore, time-lapse incubators have permitted improved morphology assessment by adding embryo kinetic events (morphokinetics) in a more stable environment during embryo development (*Kaser and Racowsky, 2014*). Although several studies have reported time-lapse algorithms to select the best embryo on day 5 (*Fishel et al., 2018*; *2020*; *Storr et al., 2018*), no conclusive evidence has yet shown that these algorithms can improve outcome (*Armstrong et al., 2019*; *Lundin and Park, 2020*).

In addition to evaluating embryo morphology, preimplantation genetic testing for aneuploidy (PGT-A) offers an advanced genetic assessment to further enhance the embryo selection process during IVF. By examining the chromosomes of embryos, PGT-A provides crucial information about their genetic health, helping fertility specialists identify embryos with the correct number of chromosomes and reduce the risk of implantation failure or miscarriage, ultimately improving the chances of a successful pregnancy (*Penzias et al., 2018*; *ESHRE PGT Consortium Steering Committee et al., 2020*).

Although some investigators have suggested that blastocyst morphology and morphokinetics (*Campbell et al., 2013*; *Capalbo et al., 2014*; *Minasi et al., 2016*) correlate with chromosomal status, it remains open to debate, and this methodology still cannot substitute PGT-A owing to its imprecision to discriminate between aneuploid and euploid embryos reliably (*Kramer et al., 2014*; *Rienzi et al., 2015*; *Zou et al., 2022*). Apart from blastocyst morphology, developmental speed is also of utmost importance. Indeed, *Bourdon et al. (2019)* showed, in a systematic review and meta-analysis, that the live birth rate (LBR) after day-5 blastocyst transfer was significantly higher than with day-6 embryos, both in fresh and frozen embryo transfer (FET) cycles (*Bourdon et al., 2019*). Nevertheless, in the current context of rapidly increasing use of vitrification, still little is known about the actual effect of the combination of vitrification day and the blastocyst quality on clinical outcomes. Only one study using untested embryos recently reported higher LBRs with good-quality day-5 embryos compared with day-6 blastocysts (*Ji et al., 2021*). Several other parameters associated with the LBR in SVBT cycles have been previously reported, including women's age, body mass index (BMI) or number of oocytes retrieved. Although these parameters can help assessing the likelihood of a live birth (*Drakopoulos et al., 2016*; *Boynukalin et al., 2020*; *Xiong et al., 2020*), they do not suffice to identify the embryo with the best prognosis for transfer.

The use of predictive modelling in medicine has recently gained interest (*Toma and Wei, 2023*), especially in the context of assisted reproduction where it can be used to predict the likelihood of implantation and live birth, hence improving patients' treatment and counselling (*Lin and Sun, 2018*). Although some prediction models of live birth after FET cycles exist (*McLernon and Bhattacharya, 2023*), most predict the cumulative live birth rate using clinical data but without using embryologic parameters, such as blastulation and vitrification day. Other studies that considered laboratory parameters while creating predictive models either included only untested embryos, focusing on maternal age as the main predictor of success (*Kato et al., 2021*), or allowed the transfer of more than one embryo without validating the model (*Coello et al., 2021*). To our knowledge, only one study reported a live birth prediction model after SVBT cycles. It did not, however, incorporate PGT-A cycles (*Xiong et al., 2022*).

The aim of the present study was to develop and validate a clinical prediction model of live birth, including blastocyst morphology and vitrification day, to help with blastocyst selection in SVBT cycles, including PGT-A tested and untested embryos.

MATERIALS AND METHODS

Study design and participants

This longitudinal monocentric retrospective cohort study was conducted in a university-based reproductive medicine unit in Spain between 2017 and 2021. It was approved by the local Institutional Review Board (Professorship in Obstetrics and Gynaecological Research of Hospital Universitari Dexeus, centre assigned to the Universidad Autónoma de Barcelona, reference: 012021012703, 27 January 2021). It was reported following the revised Transparent Reporting of Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement (*Moons et al., 2015*).

The study included all consecutive vitrified–warmed blastocysts from IVF and intracytoplasmic sperm injection cycles with autologous fresh oocytes carried out between 2017 and 2021. Only single blastocyst transfer cycles were included. Untested and euploid embryos were considered. The exclusion criteria were as follows: vitrified oocytes, oocyte or embryo donation cycles and D-grade

blastocysts. The dataset comprised the training dataset, which included all cycles between 2017 and 2020, and the validation dataset, encompassing all cycles in 2021. The indications for PGT-A were advanced maternal age, repeated implantation failure, recurrent miscarriages, severe male factor, chromosomal translocations and previous pregnancies affected by chromosomal abnormalities. Ovarian stimulation was carried out according to standard protocols described previously by *Martínez* et al. *(2016)*. Oocytes were retrieved 36 h after trigger.

Laboratory procedures

Collected oocytes were inseminated by IVF or ICSI according to sperm quality. Semen samples were assessed according to guidelines from the fifth edition of the World Health Organization Manual (*WHO, 2010*). Intracytoplasmic sperm injection and IVF were carried out according to standard procedures described previously by *Pons et al. (2019)*. Embryo culture was extended to day 7. Embryo evaluation was carried out between 114 and 118 h after insemination on day 5, and between 136 and 140 h after insemination on day 6, following the recommendations of Istanbul Consensus (*Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group Embryology, 2011*) and the Association for the Study of Reproductive Biology (ASEBIR) (*Cuevas Saiz et al., 2018*). Because of the lack of consensus about observation timings for day 7 in our centre, embryos that reached the blastocyst stage between 144 and 170 h after insemination were considered day-7 blastocysts. Blastocyst assessment was conducted according to the ASEBIR scoring system, which is based on the degree of blastocoel expansion, the trophectoderm and inner cell mass morphology, in a similar way to Gardner's grading (*Gardner and Schoolcraft, 1999*) (**Supplementary Table 1**). The ASEBIR scoring system, however, gives the trophectoderm quality more predictive strength. Hence, the final grade assigned to each blastocyst is composed of a single letter that, in most cases, matches the trophectoderm grade. Blastocyst grades range from A (the highest) to D (the lowest). Blastocysts scored as D were deemed unsuitable for clinical use and discarded. It is worth mentioning that, before 2021 in the ASEBIR scoring system, day 6A quality blastocyst did not exist (**Supplementary Table 1**).

Trophectoderm biopsy was carried out on days 5, 6 or 7, as described previously by *Coll* et al. *(2018)*. Whole-genome amplification was carried out using the Sureplex Amplification System (Illumina®, San Diego, CA, USA). The cytogenetic analysis of the biopsies was carried out by next-generation sequencing using the VeriSeq™PGS - MiSeq® platform (Illumina, San Diego CA, USA) following the manufacturer's protocols and guidelines. Embryos were diagnosed as euploid, aneuploid or mosaic. A mosaic anomaly was reported when the deviation from the normal copy number was 30% or over but less than 70% (*Coll et al., 2021*). Only the embryos classified as A, B or C were cryopreserved, all by the vitrification methodology described by *Kuwayama et al. (2005)*. After the warming process, blastocyst survival was evaluated immediately and the re-expansion was checked 2 ± 1 h after, at the time of embryo transfer.

Endometrial preparation, embryo transfer and follow-up

Hormonal replacement under standard protocol, described in details elsewhere, was used for endometrial preparation (*Alvarez, 2003*; *Martinez et al., 2016*). The endometrial preparation included exogenous oestrogens for 12–14 days. It was followed by a transvaginal ultrasound to assess endometrial thickness. Vaginal micronized progesterone treatment was added (day 0) and continued until the day of plasma beta-HCG test (days 14–16). One day before transfer (day 4) serum levels of oestradiol and progesterone were assessed. The embryo transfer was performed under transabdominal ultrasound guidance. Regardless of the day of the blastocyst vitrification (days 5, 6 or 7), it was always replaced after the same duration of progesterone treatment before transfer. Both oestradiol and progesterone were maintained until 10 weeks of pregnancy. The serum level of beta-HCG was checked 9–11 days after embryo transfer. The viability ultrasound scan was carried out around week 7 of gestation.

Primary outcome and predictor's selection

The primary outcome was live birth rate, defined as the proportion of cycles leading to the delivery of a live born after 22 weeks of gestation (*Zegers-Hochschild et al., 2017*).

The following clinical parameters were collected as potential predictors: patient age at vitrification and at embryo transfer (years), BMI (kg/m^2), anti-Müllerian hormone (AMH, ng/ml), antral follicle count (number) - both prior to hormonal stimulation, the total dose of gonadotrophins used (IU), days of ovarian stimulation, number of oocytes retrieved, semen quality, use of PGT-A, ASEBIR blastocyst grade and the day of vitrification (day 5, 6 or 7). They were selected based on previous systematic reviews and meta-analyses: maternal age (*Van Loendersloot et al., 2010*; *Neves et al., 2023*), AMH (*Practice Committee of the American Society for Reproductive Medicine, 2015*), number of oocytes and embryo quality (*Van Loendersloot et al., 2010*), BMI (*Purewal et al., 2019*); use of PGT-A (*Simopoulou et al., 2021*); vitrification day (*Bourdon et al., 2019*); blastocyst grade (*Pons et al., 2023*); and others such as dose of gonadotrophins based on our medical and embryological experience.

Statistical analysis

Sample size

In the present study, no formal sample size calculation was made. Instead, all available data in the database between January 2017 and December 2020 was used to maximize the power and generalizability of the results.

Although no sample size calculation was made, the study met all the requirements for developing a prediction and validation model (*Moons et al., 2015*). To avoid overestimating the model's performance, the dataset of the following year (January 2021 to December 2021) was used for the external validation, which included additional cycles fulfilling inclusion criteria. No imputation methods were used for missing data (as observed).

Descriptive statistics

For descriptive analysis of both groups (training and validation datasets), means and SD were used for continuous variables, whereas percentages and frequencies were used for categorical variables.

Bivariate analysis

First, a bivariate analysis was applied to identify the association between each variable with live birth. Comparison between groups was performed with either chi-squared or Student's t-test.

Multivariable analysis and creating predictive model

Second, a multivariable logistic regression model was applied to build the prediction model. The predictors included during modelling and chosen based on the strength of their unadjusted association with live birth ($P < 0.05$) in bivariate analysis were as follows: number of oocytes retrieved, PGT-A, vitrification day and ASEBIR grade. Additionally, the model was forced to include mean female age at vitrification, BMI and total dose of gonadotrophins because they were considered clinically relevant. They were included in multivariable analysis despite $P > 0.05$ in bivariate analysis as it does not necessarily imply that a non-significant predictor is not important (*Moons et al., 2015*). The day-7 blastocysts, however, included initially in bivariate analysis, were excluded from the multivariable analysis and, as a consequence from the predictive model, because of the low number of cases (25/1653 [1.5%]) to avoid model overfitting (*Moons et al., 2015*). Finally, BMI, the number of oocytes retrieved and total dose of gonadotrophins were excluded from the final model because of their statistically null beta-coefficients.

Therefore, the final predictive model estimated the probability of live birth according to ASEBIR blastocyst grade and the day of vitrification adjusted by age, use of PGT-A and the interaction between age and PGT-A cycles.

To evaluate this model's performance for development and training datasets, the receiver operating characteristic curve was adopted to calculate the area under the curve (AUC) and to determine sensitivity and specificity. The calibration curve also was analysed for training and validation datasets.

All tests were two-tailed, and $P < 0.05$ was considered statistically significant. R software (R Core Team, 2019) was used for all analyses. pROC package was used.

RESULTS

Training and validation sets: participants' characteristics

Of a total of 2972 cycles of SVBT carried out between 2017 and 2020, 1319 were excluded for different reasons; therefore, 1653 cycles fulfilled the inclusion criteria and were included in the training set (FIGURE 1).

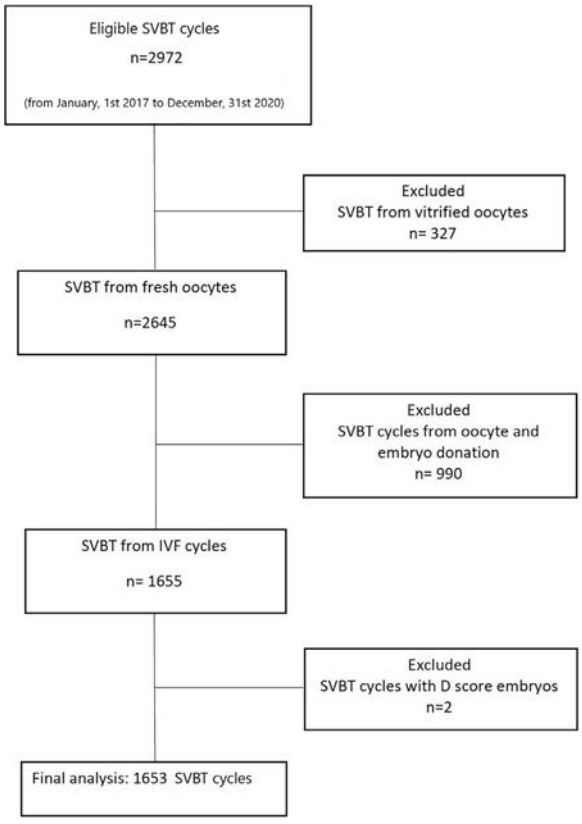


FIGURE 1 Derivation of training data set of single vitrified–warmed blastocyst transfer (SVBT).

The external validation dataset included 592 SVBT cycles carried out in 2021.

The respective baseline characteristics of both groups of patients (training and validation datasets) are presented in TABLE 1. Each variable had similar values in both sets, indicating that the model could be useful in the validation set. Comparing both sets statistically is not required, as reported for well-known and long-existing models (APACHE risk score or Framingham risk score) (*Moons et al., 2015*).

Overall live birth rate was 41.6% (688/1653) in the training dataset and 40.5% (240/592) in the validation dataset.

Variables associated with live birth rate

Bivariate analysis showed that the variables associated with live birth were number of oocytes retrieved, PGT-A use, day of vitrification and the blastocyst grade (TABLE 2). Patients with successful outcome had more oocytes retrieved than unsuccessful couples (15.2 ± 7.6 versus 14.4 ± 7.2, $P = 0.034$). Live birth rate was 48.1% when patients used PGT-A compared with 36.5% without PGT-A ($P < 0.001$). No significant difference was found in LBR between patient groups with BMI less than 30 and those with BMI 30 or over (43.2% versus 39.3%, respectively, $P = 0.547$). A higher proportion of day-5 blastocysts was found in the live birth group (76.6% [527/688]) compared with the non-live birth group (60.9% [588/965]) and a higher percentage of good-quality blastocysts (A + B) was found in the live birth group (85.7% [590/688]) versus a 70.5% (680/965) in the non-live birth group. Blastocyst categories resulting from the combination of vitrification day and ASEBIR grade also resulted in significant differences in LBR, with the live birth group had a higher proportion of the top categories (day 5A and day 5B), 68.5% (471/688) versus 49.6% (479/965) in the non- live birth group, and with a higher proportion of the lowest (day 6C) in the non-live birth group, 17.1% (165/965) versus 6% (41/688) in the live birth group.

Live birth prediction model

The prediction model for live birth was based on multivariable logistic regression. Predictors included initially in the model were selected according to the results of bivariate analysis ($P < 0.05$) and their clinical relevance (regardless of the *P*-

TABLE 1 POTENTIAL PREDICTORS IN TRAINING AND VALIDATION DATASETS

Potential predictors		Training dataset (*n* = 1653)	Validation dataset (*n* = 592)
Female age at vitrification, years		36.4 ± 3.85 (0)	36.0 ± 3.87 (1)
Female age at transfer, years		36.8 ± 3.84 (0)	37.0 ± 3.86 (0)
Body mass index, kg/m^2,[a]	<30	1134 (92.7)	326 (90.3)
	≥30	89 (7.3)	35 (9.7)
Serum AMH level, ng/ml		2.95 ± 2.56 (971)	2.91 ± 2.56 (454)
Antral follicle count		14.6 ± 8.09 (370)	15.0 ± 8.41 (200)
Total dose of gonadotrophins, IU		2378 ± 1018	2383 ± 1030
Stimulation days		10.8 ± 1.44 (69)	11.2 ± 3.6
Oocytes retrieved		14.8 ± 7.37	13.8 ± 6.9
Normal spermatozoa (WHO 5th manual), *n* (%)		1148 (69.4)	415 (70.1)
PGT-A cycles, *n* (%)		729 (44.1)	231 (39.0)
Day 5, *n* (%)		1115 (67.5)	392 (66.2)
Day 6, *n* (%)		513 (31.0)	190 (32.1)
Day 7, *n* (%)		25 (1.51)	10(1.7)
ASEBIR Score A, *n* (%)		189 (11.4)	73(12.3)
ASEBIR Score B, *n* (%)		1081 (65.4)	368(62.2)
ASEBIR Score C, *n* (%)		383 (23.2)	151 (25.5)

Results are presented as mean ± SD or *n* (%) when appropriate. Missing data are also shown in parenthesis.
[a] A total of 430 and 231 cases were missing in training and validation dataset groups, respectively.
AMH, anti-Müllerian hormone; ASEBIR, Association for the Study of Reproductive Biology
PGT-A, preimplantation genetic testing for aneuploidies; WHO, World Health Organization.

value). The predictive variables, left in the final model after removing irrelevant predictors (with null β-coefficient), are presented in TABLE 3.

Predictive probability of live birth was counted for each combination of blastocyst morphology and vitrification day adjusted by other predictors (mentioned in the Statistical analysis section). The model considered SVBT of untested embryo and day 5A as a reference groups. The latter demonstrated that, when the category was lower, live birth rate decreased: day 5B (OR 0.59, 95% CI 0.43 to 0.83), day 5C (OR 0.39, 95% CI 0.25 t o0.61), day 6B (OR 0.37, 95% CI 0.25to 0.55) and day 6C (OR 0.17, 95% CI 0.11 to 0.27).

The prediction model can be applied using the equation that follows, which was created with the values obtained by the multivariable analysis (TABLE 3):

Predicted probability of live birth = exp(w)/[1+exp(w)], where w is 3.62-0.10*(age)-5.21*(PGT-A) + 0.16*(age*PGT-A)-0.51*(D5-B)-0.94*(D5-C)-0.99*(D6-B)-1.77*(D6-C).

Finally, to estimate the exact probabilities for each group (day 5A, day 5B, day 5C, day 6B and day 6C), the mean age was set as a reference value. The PGT-A predictor was considered as a categorical variable and the predictive probabilities were calculated separately according to this variable (FIGURE 2).

Model performance

The model's original regression used for predicting live birth in training dataset resulted in an area under the curve (AUC) of 0.66 (95% CI 0.64 to 0.69) with a sensitivity of 56.0% (95% CI 52.2% to 59.8%) and a specificity of 68.5% (95% CI 65.5% to 71.3%). Calibration plot was constructed (FIGURE 3a and FIGURE 3c). Finally, an external validation was carried out on the 2021 dataset. The AUC was 0.65 (95% CI 0.61 to 0.70) with a sensitivity of 76.6% (95% CI 71.6% to 81.6%) and a specificity of 48% (95% CI 42.7% to 48%), similar to the results of training dataset (FIGURE 3b and FIGURE 3d).

DISCUSSION

To the best of our knowledge, this is the first study reporting the development and validation of a model based on pre-vitrification embryo parameters that aims to predict LBR following a single embryo transfer of either tested (PGT-A) or untested blastocyst in autologous SVBT cycles. The created multivariable prediction model shows that female age and embryo characteristics, such as overall quality and blastulation day, are associated with the LBR in SVBT cycles. Interestingly, the stratification observed for embryo vitrification day and quality (from day-5A blastocysts to day-6C blastocysts) applies to both tested (PGT-A) and untested embryos. The AUC of this predictive model demonstrated moderate discrimination (0.64), aligning well with the calibration plot, and remained consistent after external validation (0.65). Undoubtedly, incorporating additional potential predictors, such as post-warm blastocyst expansion grade, endometrial thickness, or the score of embryo transfer difficulty, could enhance the sensitivity and specificity of the model. Introducing these parameters was beyond the scope of this study, which focused on creating and validating a predictive model during the treatment phase before their use.

Some recent studies specifically constructed and explored the use of LBR prediction models after embryo transfer, although they differ from ours in some aspects. *Awadalla* et al. *(2021)* validated their model for predicting singleton, twin and total live birth rates after IVF with the transfer of one or more untested embryos in the context of both fresh and frozen embryo transfer cycles (*Awadalla et al., 2021*). *Chen* et al. *(2022)* explored the predictors for cumulative LBR and developed a prediction model based on embryo characteristics including ploidy status and an artificial intelligence algorithm for embryo selection (*Chen et al., 2022*). *Gao* et al. *(2021)* analysed only cycles with untested embryos and included heterogeneous data (IVF cycles, fresh and frozen embryo transfers, transfer of one or more embryos), but they did not consider the blastulation and vitrification day as predictors. The only other published prediction model providing guidance for embryo selection in SVBT cycles reported a similar AUC to ours (0.67) (*Xiong et al., 2022*). These investigators, however, did not consider embryos tested by PGT-A (euploid). Contrarily, they included two additional variables in their model, i.e. endometrial thickness and day-3 morphology. We did not include endometrial thickness as this potential predictor did not affect the the choice of the specific embryo for transfer.

TABLE 2 BIVARIATE ANALYSIS TO ASSESS THE ASSOCIATION BETWEEN POTENTIAL PREDICTORS WITH LIVE BIRTH IN TRAINING DATASET

Potential predictors		Live birth (*n* = 688)	No live birth (*n* = 965)	*P*-value
Mean female age at vitrification, years		36.3 ± 4.1	36.5 ± 3.7	0.408
Mean female age at transfer, years		36.7 ± 4.0	36.9 ± 3.7	0.271
Body mass index, kg/m^2[a]	<30	490	644	0.547
	≥30	35	54	
Serum AMH level, ng/ml		3.1 ± 2.6	2.8 ± 2.5	0.204
Antral follicle count		14.7 ± 8.1	14.5 ± 8.1	0.708
Total dose of gonadotrophins, IU		2385 ± 1052	2373 ± 993	0.803
Stimulation days		10.9 ± 1.4	10.8 ± 1.4	0.428
Oocytes retrieved		15.2 ± 7.6	14.4 ± 7.2	0.034
Normal sperm parameters according to WHO, *n* (%)		488 (42.5)	660 (57.5)	0.294
Abnormal sperm parameters according to WHO, *n* (%)		200 (39.6)	305 (60.4)	
PGT-A cycles, *n* (%)	Yes	351 (48.1)	378 (51.9)	<0.001
	No	337 (36.5)	587 (63.5)	
Vitrification day	5	527 (47.3)	588 (52.7)	<0.001
	6	159 (31.0)	354 (69.0)	
	7	2 (8.0)	23 (92.0)	
ASEBIR grade	A	106(56.1)	83(43.9)	<0.001
	B	484 (44.8)	597 (55.2)	
	C	98 (25.6)	285 (74.4)	
Blastocyst categories (vitrification day and ASEBIR grade)	5A	106 (56.1)	83 (43.9)	<0.001
	5B	365 (48.0)	396 (52.0)	
	5C	56 (33.9)	109 (66.1)	
	6B	118 (38.4)	189 (61.6)	
	6C	41 (19.9)	165 (80.1)	
	7B	1 (7.7)	12 (92.3)	
	7C	1 (8.3)	11 (91.7)	

[a] A total of 167 and 267 cases missing in the groups of live birth and no live birth, respectively.
Results are presented as mean ± SD or *n* (%) when appropriate. The association between each potential predictor and live birth was assessed and its statistical significance is represented by the corresponding *P*-value.
AMH, anti-Müllerian hormone; ASEBIR, Association for the Study of Reproductive Biology; PGT-A, preimplantation genetic testing for aneuploidies; WHO, World Health Organization.

We also did not include day-3 morphology because it has been shown to have no additional predictive value over day-5 morphology for live birth (*Pons et al., 2023*). The other predictors included in our study were carefully selected based on the evidence (*Van Loendersloot et al., 2010*) and their potential clinical usefulness. In addition to embryo ploidy, blastulation day, and blastocyst morphology, among the remaining parameters, only the age of women was associated with the LBR in the model. Expectedly, this association was only observed in the group of untested embryos.

Several studies examined the association between embryo characteristics and clinical outcome after FET cycles (*Ahlstrom et al., 2013*; *Yang et al., 2016*; *Irani et al., 2017*; *Cimadomo et al., 2018*; *Du et al., 2018*; *Ferreux et al., 2018*; *Bourdon et al., 2019*; *Nazem et al., 2019*; *Boynukalin et al., 2020*; *Coello et al., 2021*; *Giunco et al., 2021*; *Allen et al., 2022*). It is still difficult, however, to draw firm conclusions from this abundant research, as these studies report results from heterogeneous populations (euploid and untested embryos, donated and autologous oocytes), clinical strategies (single and double embryo transfer, cleavage and blastocyst stage) and laboratory approaches for embryo quality assessment (pre-vitrification or post-warming embryo parameters). Although these studies are useful for understanding the association between prognostic factors and LBR, the relevance of comparing them with our prediction model of success after SVBT is overall limited as they were not designed with this objective in mind. In the present study, day-5 blastocysts were associated with a significantly higher LBR

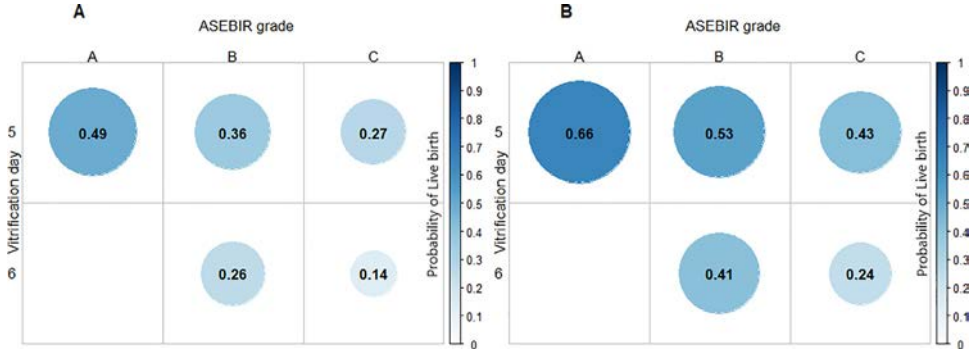


FIGURE 2 Live birth rate prediction model in (A) single vitrified–warmed blastocyst transfer (SVBT) cycles in patients with mean age, 36.8 years; and (B) SVBT preimplantation genetic testing for aneuploidy cycles, according to Association for the Study of Reproductive Biology (ASEBIR) grade (A, B or C) and the vitrification day (day 5 or day 6).

TABLE 3 MULTIVARIABLE LOGISTIC REGRESSION ANALYSIS FOR LIVE BIRTH, ADJUSTED FOR PREDICTOR VARIABLES

Intercept and predictors	Beta coefficient	Odds ratio	95% CI
Intercept	3.62	37.73	9.02 to 160.12
Age at vitrification	−0.10	0.90	0.87 to 0.94
SVBT-no PGT-A	Reference group		
SVBT-PGT-A	−5.21	0.01	0.005 to 0.05
Age*SVBT-PGT-A	0.16	1.18	1.11 to 1.25
5A category	Reference group		
5B Category	−0.51	0.59	0.43 to 0.83
5C Category	−0.94	0.39	0.25 to 0.61
6B Category	−0.99	0.37	0.25 to 0.55
6C Category	−1.77	0.17	0.11 to 0.27
AUC (95% CI)	0.66 (0.64 to 0.69)		

In the initial model, body mass index, the number of oocytes retrieved and the total dose of gonadotrophins were included. These variables, however, were excluded from the final model presented in the table because of their statistically null β-coefficients.

Age*SVBT-PGT-A, interaction between age and SVBT-PGT-A; AUC, area under the curve; PGT-A, preimplantation genetic testing for aneuploidies; SVBT, single vitrified–warmed blastocyst transfer; 5A-6C category: blastocyst score according to Association for the Study of Reproductive Biology (ASEBIR) scoring system.

than day-6 blastocysts. This is in line with the recent systematic review and meta-analysis of *Bourdon* et al. *(2019)*, showing that day-5 blastocyst transfers result in significantly higher LBR than day-6 ones (*Bourdon et al., 2019*). This could be attributed to impaired embryo metabolism (*Taylor et al., 2014*; *Minasi et al., 2016*), increased incidence of abnormal spindle, inappropriate oocyte cytoplasmatic maturation, or both (*Agca and Agca, 2014*; *Yanez et al., 2016*). According to the present model, any blastocyst formed on day 5 (grade A, B or C) yields higher LBR than even the best-quality day-6 blastocysts. Interestingly, this happened to be true for both euploid and untested embryos. Overall, these results are in concordance with the study published by *Xiong et al. (2022)*. On the other hand, the retrospective study published by *Ji* et al. *(2021)* also concluded that the embryo developmental stage and the morphological grade were useful predictors of LBR but only for untested FET cycles (*Ji et al., 2021*). According to the investigators, once a blastocyst is found to be euploid, its morphological quality is no longer associated with its implantation potential. It is important to highlight that the morphological classification in that study included only two grades (good or low quality). In contrast, the present study revealed the same ranking to select the best embryo for both untested and tested blastocysts, ranging from day 5A to day 6C (day 5A > day 5B > day 5C > day 6B > day 6C). Therefore, morphology is equally significant for euploid and untested blastocysts.

Regarding the clinical characteristics included in the model, the BMI, AMH, AFC and dose of gonadotrophins were considered potential predictors before modelling; while not considered reliable predictors of live birth in the context of FET cycles (*Li et al., 2014*; *Munch et al., 2017*; *Peigné et al., 2023*), they were described as valid markers of ovarian

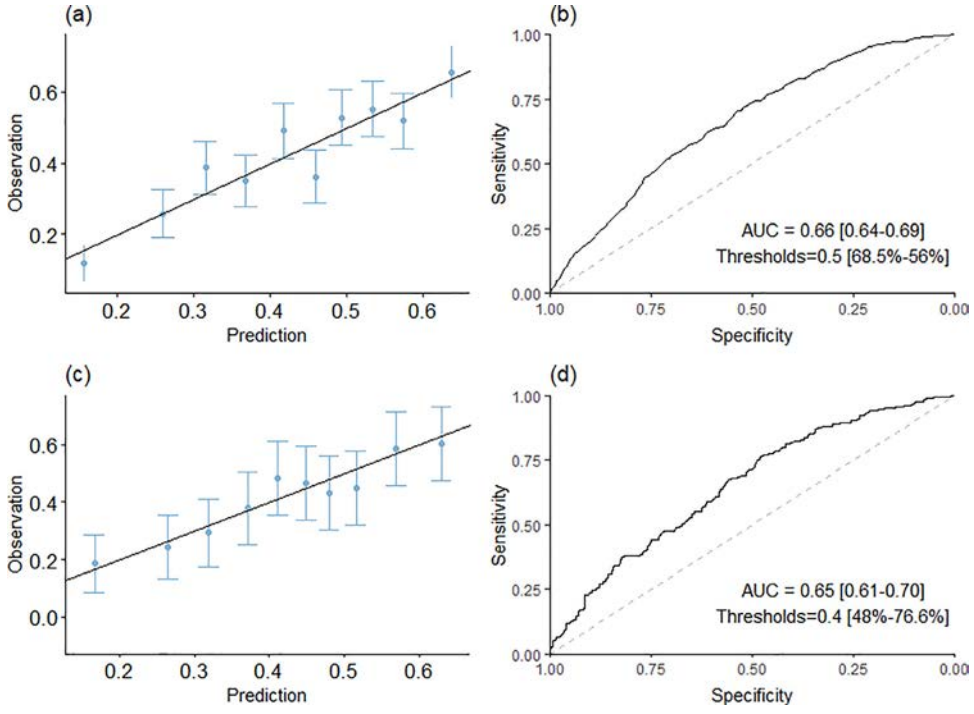


FIGURE 3 Calibration plots of the model with distribution of observed and predicted live birth rate in the training (a) and validation (c) datasets. Receiver operating curve analysis of the model's performance in the training (b) and validation (d) datasets. AUC, area under the curve.

response to stimulation (*Broekmans et al., 2006*; *Fauser et al., 2008*; *Nardo et al., 2009*). The other potential predictor that we have considered, the total number of retrieved oocytes, was found to be relevant by other authors (*Xiong et al., 2022*) and statistically significant in our unadjusted analysis. Nevertheless, in the present study, none of the aforementioned parameters was found to be statistically significant for the prediction of live birth during modelling and, as a result, they were ultimately excluded from the final model.

While the insemination technique was not considered as a potential predictor in our study because of established evidence suggesting similar live birth rates for both IVF and ICSI in cases of non-male factor infertility (*Dang et al., 2021*), our focus shifted to exploring male factor and sperm quality. We observed a comparable distribution of normal and abnormal semen results in both successful and unsuccessful groups, leading us to dismiss this parameter as a potential predictor. It is noteworthy that in an unselected group of patients, such as those who have not experienced recurrent miscarriage or implantation failure, as seen in the present study, ICSI is believed to overcome even severe male factor (*Lara-Cerrillo et al., 2021*). Nevertheless, emerging evidence suggests that nuanced analysing of specific semen parameters, such as sperm motility in IVF or morphology in ICSI cycles, may offer insights into predicting live birth outcomes (*Villani et al., 2022*).

The present study has a number of strengths. First, we included a relatively large number of cycles. Second, we included a homogeneous population of cycles only with vitrification and a single blastocyst transfer policy that allowed correlating embryo characteristics with the clinical outcome without correcting for the number of embryos transferred or the freezing method. Finally, we evaluated the performance of the model and validated it externally to determine its reproducibility and generalizability to new and different patients following Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis and Diagnosis (TRIPOD) (*Moons et al., 2015*). We, however, acknowledge that the main weakness of this study is its retrospective character, advocating for its prospective validation.

In conclusion, the pre-vitrification morphology grade and the vitrification day are associated with the live birth rate in SVBT cycles with either untested or euploid blastocysts. At this juncture, this model serves as a valuable guide, providing patients with comprehensive information not only about the quality of their embryos but also regarding their prognosis. Additionally, it aids embryologists in the prudent selection of the blastocyst for warming, thereby circumventing subjective biases. Although the developed model still has moderate value, future approximations may help increase the number of live births.

DATA AVAILABILITY

The authors do not have permission to share data.

ACKNOWLEDGMENTS

This work was carried out under the auspices of the Càtedra d'Investigació en Obstetrícia I Ginecologia of the Universitat Autónoma de Barcelona.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.rbmo.2024.103890.

REFERENCES

Adamson, G.D., Dyer, S., Chambers, G., Ishihara, O., De Mouzon, J., Kupka, M., Banker, M., Zegers-Hochschild, F., 2022. O-151 ICMART preliminary world report 2018. Human Reproduction 37,. https://doi.org/10.1093/humrep/deac105.057 deac105.057.

Agca, C., Agca, Y., 2014. Molecular and ultrastuctural changes of rat pre-implantation embryos during two-cell developmental arrest. J Assist Reprod Genet 31, 767–780. https://doi.org/10.1007/s10815-014-0213-4.

Ahlstrom, A., Westin, C., Wikland, M., Hardarson, T., 2013. Prediction of live birth in frozen-thawed single blastocyst transfer cycles by pre-freeze and post-thaw morphology. Human Reproduction 28, 1199–1209. https://doi.org/10.1093/humrep/det054.

Allen, M., Hale, L., Lantsberg, D., Kieu, V., Stevens, J., Stern, C., Gardner, D.K., Mizrachi, Y., 2022. Post-warming embryo morphology is associated with live birth: a cohort study of single vitrified-warmed blastocyst transfer cycles. J Assist Reprod Genet 39, 417–425. https://doi.org/10.1007/s10815-021-02390-z.

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, Balaban, B., Brison, D., Calderon, G., Catt, J., Conaghan, J., Cowan, L., Ebner, T., Gardner, D., Hardarson, T., Lundin, K., Cristina Magli, M., Mortimer, D., Mortimer, S., Munne, S., Royere, D., Scott, L., Smitz, J., Thornhill, A., van Blerkom, J., Van den Abbeel, E., 2011. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Human Reproduction 26, 1270–1283. https://doi.org/10.1093/humrep/der037.

Alpha Scientists in Reproductive Medicine, ESHRE Special Interest Group Embryology, 2011. Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Reproductive BioMedicine Online 22, 632–646. https://doi.org/10.1016/j.rbmo.2011.02.001.

Alvarez, C., 2003. Biological variation of seminal parameters in healthy subjects. Human Reproduction 18, 2082–2088. https://doi.org/10.1093/humrep/deg430.

Armstrong, S., Bhide, P., Jordan, V., Pacey, A., Marjoribanks, J., Farquhar, C., 2019. Time-lapse systems for embryo incubation and assessment in assisted reproduction. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.CD011320.pub4.

Awadalla, M.S., Bendikson, K.A., Ho, J.R., McGinnis, L.K., Ahmady, A., 2021. A validated model for predicting live birth after embryo transfer. Sci Rep 11, 10800. https://doi.org/10.1038/s41598-021-90254-y.

Bourdon, M., Pocate-Cheriet, K., Finet de Bantel, A., Grzegorczyk-Martin, V., Amar Hoffet, A., Arbo, E., Poulain, M., Santulli, P., 2019. Day 5 versus Day 6 blastocyst transfers: a systematic review and meta-analysis of clinical outcomes. Human Reproduction 34, 1948–1964. https://doi.org/10.1093/humrep/dez163.

Boynukalin, F.K., Gultomruk, M., Cavkaytar, S., Turgut, E., Findikli, N., Serdarogullari, M., Coban, O., Yarkiner, Z., Rubio, C., Bahceci, M., 2020. Parameters impacting the live birth rate per transfer after frozen single euploid blastocyst transfer. PLoS ONE 15, e0227619. https://doi.org/10.1371/journal.pone.0227619.

Broekmans, F.J., Kwee, J., Hendriks, D.J., Mol, B.W., Lambalk, C.B., 2006. A systematic review of tests

predicting ovarian reserve and IVF outcome. Human Reproduction Update 12, 685–718. https://doi.org/10.1093/humupd/dml034.
Campbell, A., Fishel, S., Bowman, N., Duffy, S., Sedler, M., Hickman, C.F.L., 2013. Modelling a risk classification of aneuploidy in human embryos using non-invasive morphokinetics. Reproductive BioMedicine Online 26, 477–485. https://doi.org/10.1016/j.rbmo.2013.02.006.
Capalbo, A., Rienzi, L., Cimadomo, D., Maggiulli, R., Elliott, T., Wright, G., Nagy, Z.P., Ubaldi, F.M., 2014. Correlation between standard blastocyst morphology, euploidy and implantation: an observational study in two centers involving 956 screened blastocysts. Human Reproduction 29, 1173–1181. https://doi.org/10.1093/humrep/deu033.
Chen, L., Li, W., Liu, Y., Peng, Z., Cai, L., Zhang, N., Xu, J., Wang, L., Teng, X., Yao, Y., Zou, Y., Ma, M., Liu, J., Lu, S., Sun, H., Yao, B., 2022. Non-invasive embryo selection strategy for clinical IVF to avoid wastage of potentially competent embryos. Reproductive BioMedicine Online 45, 26–34. https://doi.org/10.1016/j.rbmo.2022.03.006.
Cimadomo, D., Capalbo, A., Levi-Setti, P.E., Soscia, D., Orlando, G., Albani, E., Parini, V., Stoppa, M., Dovere, L., Tacconi, L., Ievoli, E., Maggiulli, R., Ubaldi, F.M., Rienzi, L., 2018. Associations of blastocyst features, trophectoderm biopsy and other laboratory practice with post-warming behavior and implantation. Human Reproduction 33, 1992–2001. https://doi.org/10.1093/humrep/dey291.
Coello, A., Meseguer, M., Galán, A., Alegre, L., Remohí, J., Cobo, A., 2017. Analysis of the morphological dynamics of blastocysts after vitrification/warming: defining new predictive variables of implantation. Fertility and Sterility 108, 659–666. https://doi.org/10.1016/j.fertnstert.2017.07.1157 e4.
Coello, A., Nohales, M., Meseguer, M., de los Santos, M.J., Remohí, J., Cobo, A., 2021. Prediction of embryo survival and live birth rates after cryotransfers of vitrified blastocysts. Reproductive BioMedicine Online 42, 881–891. https://doi.org/10.1016/j.rbmo.2021.02.013.
Coll, L., Parriego, M., Boada, M., Devesa, M., Arroyo, G., Rodríguez, I., Coroleu, B., Vidal, F., Veiga, A., 2018. Transition from blastomere to trophectoderm biopsy: comparing two preimplantation genetic testing for aneuploidies strategies. Zygote 26, 191–198. https://doi.org/10.1017/S0967199418000084.
Coll, L., Parriego, M., Mateo, S., García-Monclús, S., Rodríguez, I., Boada, M., Coroleu, B., Polyzos, N.P., Vidal, F., Veiga, A., 2021. Prevalence, types and possible factors influencing mosaicism in IVF blastocysts: results from a single setting. Reproductive BioMedicine Online 42, 55–65. https://doi.org/10.1016/j.rbmo.2020.09.025.
Cuevas Saiz, I., Carme Pons Gatell, M., Vargas, M.C., Delgado Mendive, A., Rives Enedáguila, N., Moragas Solanes, M., Carrasco Canal, B., Teruel López, J., Busquets Bonet, A., Hurtado de Mendoza Acosta, M.V., 2018. The Embryology Interest Group: updating ASEBIR's morphological scoring system for early embryos, morulae and blastocysts. Medicina Reproductiva y Embriología Clínica 5, 42–54. https://doi.org/10.1016/j.medre.2017.11.002.
Dang, V.Q., Vuong, L.N., Luu, T.M., Pham, T.D., Ho, T.M., Ha, A.N., Truong, B.T., Phan, A.K., Nguyen, D.P., Pham, T.N., Pham, Q.T., Wang, R., Norman, R.J., Mol, B.W., 2021. Intracytoplasmic sperm injection versus conventional in-vitro fertilisation in couples with infertility in whom the male partner has normal total sperm count and motility: an open-label, randomised controlled trial. The Lancet 397, 1554–1563. https://doi.org/10.1016/S0140-6736(21)00535-3.
Drakopoulos, P., Blockeel, C., Stoop, D., Camus, M., De Vos, M., Tournaye, H., Polyzos, N.P., 2016. Conventional ovarian stimulation and single embryo transfer for IVF/ICSI. How many oocytes do we need to maximize cumulative live birth rates after utilization of all fresh and frozen embryos? Hum. Reprod.. https://doi.org/10.1093/humrep/dev316 dev316.
Du, T., Wang, Y., Fan, Y., Zhang, S., Yan, Z., Yu, W., Xi, Q., Chen, Q., Mol, B.W., Lyu, Q., Kuang, Y., 2018. Fertility and neonatal outcomes of embryos achieving blastulation on Day 7: are they of clinical value? Human Reproduction 33, 1038–1051. https://doi.org/10.1093/humrep/dey092.
ESHRE PGT Consortium Steering Committee, Carvalho, F., Coonen, E., Goossens, V., Kokkali, G., Rubio, C., Meijer-Hoogeveen, M., Moutou, C., Vermeulen, N., De Rycke, M., 2020. ESHRE PGT Consortium good practice recommendations for the organisation of PGT†. Human Reproduction Open 2020,. https://doi.org/10.1093/hropen/hoaa021 hoaa021.
European IVF Monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Gliozheni, O., Hambartsoumian, E., Strohmer, H., Petrovskaya, E., Tishkevich, O., De Neubourg, D., Bogaerts, K., Balic, D., Antonova, I., Cvetkova, E., Rezabek, K., Kirk, J., Sõritsa, D., Gissler, M., Pelkonen, S., Mansouri, I., De Mouzon, J., Tandler-Schneider, A., Kimmel, M., Vrachnis, N., Urbancsek, J., Kosztolanyi, G., Bjorgvinsson, H., Wingfield, M., Leyden, J., Scaravelli, G., De Luca, R., Lokshin, V., Karibayeva, S., Agloniete, V., Bausyte, R., Masliukaite, I., Schilling, C., Calleja-Agius, J., Moshin, V., Simic, T.M., Vukicevic, D., Smeenk, J.M.J., Petanovski, Z., Romundstad, L.B., Janicka, A., Calhaz-Jorge, C., Guimaraes, J.M.M., E Silva, P.D., Korsak, V., Vidakovic, S., Marsik, L., Kovacic, B., Saiz, I.C., Mondéjar, F.P., Bergh, C., Toitot, S., Schneider, M., Isikoglu, M., Balaban, B., Gryshchenko, M., Bridges, E., Ewans, A., Smeenk, J., Wyns, C., De Geyter, C., Kupka, M., Bergh, C., Cuevas Saiz, I., De Neubourg, D., Rezabek, K., Tandler-Schneider, A., Rugescu, I., Goossens, V., 2023. ART in Europe, 2019: results generated from European registries by ESHRE. Human Reproduction dead197. https://doi.org/10.1093/humrep/dead197.
Fauser, B.C.J.M., Diedrich, K., Devroey, P., 2008. Predictors of ovarian response: progress towards individualized treatment in ovulation induction and ovarian stimulation. Human Reproduction Update 14, 1–14. https://doi.org/10.1093/humupd/dmm034.
Ferreux, L., Bourdon, M., Sallem, A., Santulli, P., Barraud-Lange, V., Le Foll, N., Maignien, C., Chapron, C., de Ziegler, D., Wolf, J.-P., Pocate-Cheriet, K., 2018. Live birth rate following frozen–thawed blastocyst transfer is higher with blastocysts expanded on Day 5 than on Day 6. Human Reproduction 33, 390–398. https://doi.org/10.1093/humrep/dey004.
Fishel, S., Campbell, A., Foad, F., Davies, L., Best, L., Davis, N., Smith, R., Duffy, S., Wheat, S., Montgomery, S., Wachter, A., Beccles, A., 2020. Evolution of embryo selection for IVF from subjective morphology assessment to objective time-lapse algorithms improves chance of live birth. Reproductive BioMedicine Online 40, 61–70. https://doi.org/10.1016/j.rbmo.2019.10.005.
Fishel, S., Campbell, A., Montgomery, S., Smith, Rachel, Nice, L., Duffy, S., Jenner, L., Berrisford, K., Kellam, L., Smith, Rob, Foad, F., Beccles, A., 2018. Time-lapse imaging algorithms rank human preimplantation embryos according to the probability of live birth. Reproductive BioMedicine Online 37, 304–313. https://doi.org/10.1016/j.rbmo.2018.05.016.
Gao, H., Liu, D., Li, Y., Wu, X., Tan, H., 2021. Early prediction of live birth for assisted reproductive technology patients: a convenient and practical prediction model. Sci Rep 11, 331. https://doi.org/10.1038/s41598-020-79308-9.
Gardner, D.K., Schoolcraft, W.B., 1999. Culture and transfer of human blastocysts: Current Opinion in Obstetrics and Gynaecology 11, 307–311. https://doi.org/10.1097/00001703-199906000-00013
Giunco, H., Connerney, M., Boylan, C., Koelper, N., Mersereau, J., Berger, D.S., 2021. Embryo re-expansion does not affect clinical pregnancy rates in frozen embryo transfer cycles: a retrospective study. J Assist Reprod Genet 38, 2933–2939. https://doi.org/10.1007/s10815-021-02319-6.
Irani, M., Reichman, D., Robles, A., Melnick, A., Davis, O., Zaninovic, N., Xu, K., Rosenwaks, Z., 2017. Morphologic grading of euploid blastocysts influences implantation and ongoing pregnancy rates. Fertility and Sterility 107, 664–670. https://doi.org/10.1016/j.fertnstert.2016.11.012.
Irani, M., Zaninovic, N., Rosenwaks, Z., Xu, K., 2019. Does maternal age at retrieval influence the implantation potential of euploid blastocysts? American Journal of Obstetrics and Gynecology 220,. https://doi.org/10.1016/j.ajog.2018.11.1103 379.e1-379.e7.
Ji, H., Zhou, Y., Cao, S., Zhang, J., Ling, X., Zhao, C., Shen, R., 2021. Effect of Embryo Developmental Stage, Morphological Grading, and Ploidy Status on Live Birth Rate in Frozen Cycles of Single Blastocyst Transfer. Reprod. Sci. 28, 1079–1091. https://doi.org/10.1007/s43032-020-00381-6.
Kaser, D.J., Racowsky, C., 2014. Clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring: a systematic review. Human Reproduction Update 20, 617–631. https://doi.org/10.1093/humupd/dmu023.
Kato, K., Ueno, S., Berntsen, J., Ito, M., Shimazaki, K., Uchiyama, K., Okimura, T., 2021. Comparing prediction of ongoing pregnancy and live birth outcomes in patients with advanced and younger maternal age patients using KIDScore™ day 5: a large-cohort retrospective study with single vitrified-warmed blastocyst transfer. Reprod Biol Endocrinol 19, 98. https://doi.org/10.1186/s12958-021-00767-4.
Kramer, Y.G., Kofinas, J.D., Melzer, K., Noyes, N., McCaffrey, C., Buldo-Licciardi, J., McCulloh, D.H., Grifo, J.A., 2014. Assessing morphokinetic parameters via time lapse microscopy (TLM) to predict euploidy: are aneuploidy risk classification models universal? J Assist Reprod Genet 31, 1231–1242. https://doi.org/10.1007/s10815-014-0285-1.
Kuwayama, M., Vajta, G., Kato, O., Leibo, S.P., 2005. Highly efficient vitrification method for cryopreservation of human oocytes. Reproductive BioMedicine Online 11, 300–308. https://doi.org/10.1016/S1472-6483(10)60837-1.
Lara-Cerrillo, S., Ribas-Maynou, J., Rosado-Iglesias, C., Lacruz-Ruiz, T., Benet, J.,

García-Peiró, A., 2021. Sperm selection during ICSI treatments reduces single- but not double-strand DNA break values compared to the semen sample. J Assist Reprod Genet 38, 1187–1196. https://doi.org/10.1007/s10815-021-02129-w.

Li, H.W.R., Lee, V.C.Y., Lau, E.Y.L., Yeung, W.S.B., Ho, P.-C., Ng, E.H.Y., 2014. Role of baseline antral follicle count and anti-Mullerian hormone in the index stimulation cycle of IVF treatment in predicting outcome of subsequent frozen-thawed embryo transfers. Gynecological Endocrinology 30, 490–493. https://doi.org/10.3109/09513590.2014.899572.

Lin, J., Sun, X.-X., 2018. Predictive Modeling in Reproductive Medicine. Reproductive and Developmental Medicine 2, 224–229. https://doi.org/10.4103/2096-2924.249888.

Lundin, K., Park, H., 2020. Time-lapse technology for embryo culture and selection. Upsala Journal of Medical Sciences 125, 77–84. https://doi.org/10.1080/03009734.2020.1728444.

Martinez, F., Rodriguez, I., Devesa, M., Buxaderas, R., Gómez, M.J., Coroleu, B., 2016. Should progesterone on the human chorionic gonadotropin day still be measured? Fertility and Sterility 105, 86–92. https://doi.org/10.1016/j.fertnstert.2015.09.008.

McLernon, D.J., Bhattacharya, S., 2023. Quality of clinical prediction models in in vitro fertilisation: Which covariates are really important to predict cumulative live birth and which models are best? Best Practice & Research Clinical Obstetrics & Gynaecology 86, 102309. https://doi.org/10.1016/j.bpobgyn.2022.102309.

Minasi, M.G., Colasante, A., Riccio, T., Ruberti, A., Casciani, V., Scarselli, F., Spinella, F., Fiorentino, F., Varricchio, M.T., Greco, E., 2016. Correlation between aneuploidy, standard morphology evaluation and morphokinetic development in 1730 biopsied blastocysts: a consecutive case series study. Hum. Reprod. 31, 2245–2254. https://doi.org/10.1093/humrep/dew183.

Moons, K.G.M., Altman, D.G., Reitsma, J.B., Ioannidis, J.P.A., Macaskill, P., Steyerberg, E.W., Vickers, A.J., Ransohoff, D.F., Collins, G.S., 2015. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 162, W1–W73. https://doi.org/10.7326/M14-0698.

Munch, E.M., Sparks, A.E., Zimmerman, M.B., Van Voorhis, B.J., Duran, E.H., 2017. High FSH dosing is associated with reduced live birth rate in fresh but not subsequent frozen embryo transfers. Human Reproduction 32, 1402–1409. https://doi.org/10.1093/humrep/dex094.

Nardo, L.G., Gelbaya, T.A., Wilkinson, H., Roberts, S.A., Yates, A., Pemberton, P., Laing, I., 2009. Circulating basal anti-Müllerian hormone levels as predictor of ovarian response in women undergoing ovarian stimulation for in vitro fertilization. Fertility and Sterility 92, 1586–1593. https://doi.org/10.1016/j.fertnstert.2008.08.127.

Nazem, T.G., Sekhon, L., Lee, J.A., Overbey, J., Pan, S., Duke, M., Briton-Jones, C., Whitehouse, M., Copperman, A.B., Stein, D.E., 2019. The correlation between morphology and implantation of euploid human blastocysts. Reproductive BioMedicine Online 38, 169–176. https://doi.org/10.1016/j.rbmo.2018.10.007.

Neves, A.R., Garcia, S., Vuong, L.T., Blockeel, C., Arroyo, G., Spits, C., Pham, T.D., Ho, T.M., Tournaye, H., Polyzos, N.P., 2023. Association between sequence variants in the FSHR gene and reproductive outcomes following IVF in predicted normoresponders. Reproductive BioMedicine Online 46, 826–834. https://doi.org/10.1016/j.rbmo.2023.01.013.

Peigné, M., Bernard, V., Dijols, L., Creux, H., Robin, G., Hocké, C., Grynberg, M., Dewailly, D., Sonigo, C., 2023. Using serum anti-Müllerian hormone levels to predict the chance of live birth after spontaneous or assisted conception: a systematic review and meta-analysis. Human Reproduction 38, 1789–1806. https://doi.org/10.1093/humrep/dead147.

Penzias, A., Bendikson, K., Butts, S., Coutifaris, C., Falcone, T., Fossum, G., Gitlin, S., Gracia, C., Hansen, K., La Barbera, A., Mersereau, J., Odem, R., Paulson, R., Pfeifer, S., Pisarska, M., Rebar, R., Reindollar, R., Rosen, M., Sandlow, J., Vernon, M., Widra, E., 2018. The use of preimplantation genetic testing for aneuploidy (PGT-A): a committee opinion. Fertility and Sterility 109, 429–436. https://doi.org/10.1016/j.fertnstert.2018.01.002.

Pons, M.C., Carrasco, B., Parriego, M., Boada, M., González-Foruria, I., Garcia, S., Coroleu, B., Barri, P.N., Veiga, A., 2019. Deconstructing the myth of poor prognosis for fast-cleaving embryos on day 3. Is it time to change the consensus? J Assist Reprod Genet 36, 2299–2305. https://doi.org/10.1007/s10815-019-01574-y.

Pons, M.C., Carrasco, B., Rives, N., Delgado, A., Martínez-Moro, A., Martínez-Granados, L., RodriguezIgnacio, I., Cairó, O., Cuevas-Saiz, I., 2023. Predicting the likelihood of live birth: an objective and user-friendly blastocyst grading system. Reproductive BioMedicine Online S1472648323003413. https://doi.org/10.1016/j.rbmo.2023.05.015

Purewal, S., Chapman, S.C.E., Van Den Akker, O.B.A., 2019. A systematic review and meta-analysis of lifestyle and body mass index predictors of successful assisted reproductive technologies. Journal of Psychosomatic Obstetrics & Gynecology 40, 2–18. https://doi.org/10.1080/0167482X.2017.1403418.

Rienzi, L., Capalbo, A., Stoppa, M., Romano, S., Maggiulli, R., Albricci, L., Scarica, C., Farcomeni, A., Vajta, G., Ubaldi, F.M., 2015. No evidence of association between blastocyst aneuploidy and morphokinetic assessment in a selected population of poor-prognosis patients: a longitudinal cohort study. Reproductive BioMedicine Online 30, 57–66. https://doi.org/10.1016/j.rbmo.2014.09.012.

Simopoulou, M., Sfakianoudis, K., Maziotis, E., Tsioulou, P., Grigoriadis, S., Rapani, A., Giannelou, P., Asimakopoulou, M., Kokkali, G., Pantou, A., Nikolettos, K., Vlahos, N., Pantos, K., 2021. PGT-A: who and when? A systematic review and network meta-analysis of RCTs. J Assist Reprod Genet 38, 1939–1957. https://doi.org/10.1007/s10815-021-02227-9.

Storr, A., Venetis, C., Cooke, S., Kilani, S., Ledger, W., 2018. Time-lapse algorithms and morphological selection of day-5 embryos for transfer: a preclinical validation study. Fertility and Sterility 109, 276–283. https://doi.org/10.1016/j.fertnstert.2017.10.036 e3.

Taylor, T.H., Patrick, J.L., Gitlin, S.A., Michael Wilson, J., Crain, J.L., Griffin, D.K., 2014. Outcomes of blastocysts biopsied and vitrified once versus those cryopreserved twice for euploid blastocyst transfer. Reproductive BioMedicine Online 29, 59–64. https://doi.org/10.1016/j.rbmo.2014.03.001.

Practice Committee of the American Society for Reproductive Medicine, 2015. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertility and Sterility 103, e9–e17. https://doi.org/10.1016/j.fertnstert.2014.12.093.

Toma, M., Wei, O.C., 2023. Predictive Modeling in Medicine. Encyclopedia 3, 590–601. https://doi.org/10.3390/encyclopedia3020042.

Van Loendersloot, L.L., Van Wely, M., Limpens, J., Bossuyt, P.M.M., Repping, S., Van Der Veen, F., 2010. Predictive factors in in vitro fertilization (IVF): a systematic review and meta-analysis. Human Reproduction Update 16, 577–589. https://doi.org/10.1093/humupd/dmq015.

Villani, M.T., Morini, D., Spaggiari, G., Falbo, A.I., Melli, B., La Sala, G.B., Romeo, M., Simoni, M., Aguzzoli, L., Santi, D., 2022. Are sperm parameters able to predict the success of assisted reproductive technology? A retrospective analysis of over 22,000 assisted reproductive technology cycles. Andrology 10, 310–321. https://doi.org/10.1111/andr.13123.

World HealthOrganization, 2010. WHO laboratory manual for the examination and processing of human semen, 5th ed. World Health Organization, Geneva. ed.

Xiong, F., Sun, Q., Li, G., Yao, Z., Chen, P., Wan, C., Zhong, H., Zeng, Y., 2020. Association between the number of top-quality blastocysts and live births after single blastocyst transfer in the first fresh or vitrified—warmed IVF/ICSI cycle. Reproductive BioMedicine Online 40, 530–537. https://doi.org/10.1016/j.rbmo.2020.01.005.

Xiong, F., Sun, Q., Wang, S., Yao, Z., Chen, P., Wan, C., Zhong, H., Zeng, Y., 2022. A nomogram to assist blastocyst selection in vitrified-warmed embryo transfer cycles. J of Obstet and Gynaecol 48, 1816–1828. https://doi.org/10.1111/jog.15138.

Yanez, L.Z., Han, J., Behr, B.B., Pera, R.A.R., Camarillo, D.B., 2016. Human oocyte developmental potential is predicted by mechanical properties within hours after fertilization. Nat Commun 7, 10809. https://doi.org/10.1038/ncomms10809.

Yang, H., Yang, Q., Dai, S., Li, G., Jin, H., Yao, G., Sun, Y., 2016. Comparison of differences in development potentials between frozen-thawed D5 and D6 blastocysts and their relationship with pregnancy outcomes. J Assist Reprod Genet 33, 865–872. https://doi.org/10.1007/s10815-016-0712-6.

Zegers-Hochschild, F., Adamson, G.D., Dyer, S., Racowsky, C., De Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., Cooke, I.D., Simpson, J.L., Van Der Poel, S., 2017. The International Glossary on Infertility and Fertility Care, 2017. Fertility and Sterility 108, 393–406. https://doi.org/10.1016/j.fertnstert.2017.06.005.

Zou, Y., Pan, Y., Ge, N., Xu, Y., Gu, R., Li, Z., Fu, J., Gao, J., Sun, X., Sun, Y., 2022. Can the combination of time-lapse parameters and clinical features predict embryonic ploidy status or implantation? Reproductive BioMedicine Online 45, 643–651. https://doi.org/10.1016/j.rbmo.2022.06.007.

Received 29 September 2023; received in revised form 28 December 2023; accepted 7 February 2024.

RBMO

ELSEVIER

ARTICLE

Progesterone concentrations on blastocyst transfer day in modified natural cycle frozen embryo transfer cycles

BIOGRAPHY

Marte Saupstad obtained her medical degree from University of Copenhagen, Denmark, in 2019. She was then a PhD student at the Fertility Department, Copenhagen University Hospital, Rigshospitalet, managing an ongoing randomized controlled study investigating the effect of progesterone luteal phase support and timing of blastocyst transfer on reproductive outcomes in modified natural cycle frozen embryo transfer.

Marte Saupstad[1,*,#], Sara J. Bergenheim[1,#], Jeanette W. Bogstad[1], Morten R. Petersen[1], Anna Klajnbard[2], Lisbeth Prætorius[3], Nina l.C. Freiesleben[3,9], Anne L. Englund[4], Ellen C.L. Løkkegaard[5,9], Ulla B. Knudsen[6], Merete Husth[7], Birgit Alsbjerg[8], Julie E. Møller[1], Tine V. Dam[1], Julie L. Forman[10], Anja Pinborg[1,9,†], Kristine Løssl[1,9,†]

KEY MESSAGE

Literature shows that progesterone concentrations <30 nmol/l on the day of blastocyst transfer are associated with poorer reproductive outcomes in frozen embryo transfer (FET) cycles. We found no association between low progesterone concentrations and poor reproductive outcomes in modified natural cycle FET. Endometrial preparation protocols should be compared with caution.

ABSTRACT

Research question: Are serum progesterone concentrations on the day of modified natural cycle (mNC) frozen blastocyst transfer (FET) without luteal phase support (LPS) associated with clinical pregnancy rate (CPR)?

Design: Data were collected between January 2019 and October 2022 as a sub-study of an ongoing randomized controlled trial assessing pregnancy outcomes in mNC-FET. The sub-study included all women ($n = 209$) randomized to mNC-FET without LPS at the time of data extraction. Participants were aged 18–41 years, had regular menstrual cycles and underwent mNC-FET treatment with single-blastocyst transfer. Associations between the serum progesterone concentration on the day of blastocyst transfer and CPR, pregnancy rate and pregnancy loss rate (PLR) were examined between groups with low and higher

[1] The Fertility Clinic, Department of Gynaecology, Fertility and Obstetrics, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark.
[2] Department of Obstetrics and Gynaecology, The Fertility Clinic, Copenhagen University Hospital Herlev, Herlev, Denmark.
[3] Department of Obstetrics and Gynaecology, The Fertility Clinic, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark.
[4] Department of Obstetrics and Gynaecology, The Fertility Clinic, Zealand University Hospital Koege, Koege, Denmark.
[5] Department of Obstetrics and Gynaecology, Copenhagen University Hospital North Zealand, North Zealand Hospital, Hilleroed, Denmark.
[6] Department of Obstetrics and Gynaecology, The Fertility Clinic, The Regional Hospital Horsens, Horsens, Denmark.
[7] The Fertility Unit, Aalborg University Hospital, Aalborg University Hospital South, Aalborg, Denmark.
[8] The Fertility Clinic, Skive Regional Hospital, Skive, Denmark.
[9] Institute of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
[10] Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.
[#] These authors should be regarded as joint first authors.
[†] These authors should be regarded as joint last authors.

*Corresponding author. E-mail address: marte.saupstad@regionh.dk (M. Saupstad). https://doi.org/10.1016/j.rbmo.2024.103862

Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS

Blastocyst transfer
Clinical pregnancy rate
Frozen embryo transfer
Modified natural cycle
Progesterone

progesterone concentrations using the 25th and 10th percentiles as cut-offs. Multivariate logistic regression analyses were performed to adjust for potential confounding factors.

Results: Progesterone concentrations on the day of blastocyst transfer in mNC-FET without LPS ranged from 4.9 to 91.8 nmol/l, with the 25th and 10th percentiles at 29.0 nmol/l and 22.5 nmol/l, respectively. Serum progesterone concentrations did not differ between women with or without a clinical pregnancy (mean [SD] 38.5 [14.0] versus 36.8 [12.4] nmol/l; $P = 0.350$). Furthermore, the CPR, pregancy rate and PLR were similar in women with low or high progesterone concentrations when using the 25th or the 10th progesterone percentile as cut-off. Multivariate regression analyses showed no association between progesterone concentrations and CPR.

Conclusions: No association was found between progesterone concentration on the day of blastocyst transfer and pregnancy outcome in women undergoing mNC-FET without progesterone LPS.

INTRODUCTION

The advances in IVF cryopreservation techniques since the early 2000s have made it safe and efficient to vitrify and warm embryos, with a high survival rate (*Rienzi et al., 2017*). Today, frozen embryo transfer (FET) is a worthy alternative to fresh embryo transfers when the risk of ovarian hyperstimulation syndrome is significant, as outcomes after fresh and frozen embryo transfer seem to be comparable (*Stormlund et al., 2020*; *Zaat et al., 2021*). Single-embryo transfers in fresh cycles are encouraged when a good cryopreservation programme is available.

Several different protocols can be used for endometrial preparation in FET cycles. A common approach is preparation with hormone replacement therapy (HRT). As women undergoing HRT-FET do not ovulate, and therefore do not develop a corpus luteum, it is of great clinical interest whether an optimal concentration of progesterone, by means of progesterone supplementation, can be established. Several papers addressing this matter have been published, and while they differ in methodology, there seems to be an overall agreement that low progesterone concentrations may have a negative impact on pregnancy outcome (*Alsbjerg et al., 2020*; *Basnayake et al., 2018*; *Brady et al., 2014*; *Cédrin-Durnerin et al., 2019*; *Labarta et al., 2021*, *2017*; *Volovsky et al., 2020*; *Yovich et al., 2015*) and that there is a minimum progesterone concentration around 30 nmol/l below which pregnancy rates in HRT-FET are suboptimal (*Cédrin-Durnerin et al., 2019*; *Labarta et al., 2021*, *2017*).

Extensive research has shown that reproductive outcomes in natural cycles and HRT-FET cycles are comparable, or maybe slightly better in natural cycles (*Ghobara et al., 2017*; *Groenewoud et al., 2017*; *Wu et al., 2021*). In addition, the risk of hypertensive disorders of pregnancy is lower in natural cycle FET (NC-FET) compared with HRT-FET (*Busnelli et al., 2022*; *Moreno-Sepulveda et al., 2021*; *von Versen-Höynck et al., 2019*). Accordingly, NC-FET, predominantly modified natural cycle FET (mNC-FET), is the preferred choice for ovulatory women at public fertility clinics in Denmark and is commonly performed without luteal phase support (LPS), i.e. administration of exogenous progesterone.

In NC-FET, transvaginal ultrasound scans, with or without blood and/or urine analyses, are used to determine the optimal timing for embryo transfer with (mNC-FET) or without (true NC-FET [tNC-FET]) the use of an exogenous ovulation-trigger. Compared with HRT-FET, far less is known about progesterone concentration around the time of blastocyst transfer, and its impact on pregnancy rates, in tNC- and mNC-FET cycles. In 2020, Gaggiotti-Marre and colleagues ($n = 294$) explored progesterone concentrations on the day before blastocyst transfer in tNC without LPS and found that progesterone concentrations below 31.8 nmol/l were associated with lower live birth rates (LBR) (*Gaggiotti-Marre et al., 2020*). Ramezenali and co-workers ($n = 101$), on the other hand, examined progesterone concentrations on the day of cleavage-stage embryo transfer in mNC-FET without LPS and found no association between progesterone concentration and clinical pregnancy rate (CPR) (*Ramezanali et al., 2019*).

To the current authors' knowledge, this is the first study investigating whether progesterone concentrations on the day of blastocyst transfer in mNC-FET cycles are associated with treatment success. Using data from a prospectively collected dataset, derived from an ongoing randomized controlled trial (RCT) (*Saupstad et al., 2019*), they sought to investigate whether progesterone concentrations on the day of blastocyst transfer in mNC-FET without LPS were associated with CPR and whether there was a cut-off progesterone value below which reproductive outcomes are compromised.

MATERIALS AND METHODS

Study design

The data included in the present study were collected at eight public hospitals in Denmark, from January 2019 to October 2022, as part of an ongoing, multicentre RCT. The RCT was designed to examine the effect of LPS (Lutinus), and the optimal timing of blastocyst transfer, on reproductive outcomes in mNC-FET. Participants were randomized using a minimization algorithm including study site, female age, number of previous IVF/intracytoplasmic sperm injection treatments, number of FET treatments and blastocyst quality to one the following groups:

A. LPS + blastocyst transfer on day 6 following the ovulation trigger;
B. LPS + blastocyst transfer on day 7 following the ovulation trigger;
C. no LPS + blastocyst transfer on day 6 following the ovulation trigger;
D. no LPS + blastocyst transfer on day 7 following the ovulation trigger.

Only participants who did not receive exogenous progesterone in the luteal phase (groups C and D), were included in the present sub-study. A full description of the RCT study protocol and mNC-FET treatment has previously been published (*Saupstad et al., 2019*).

Study population

The present sub-study included a total of 209 women, consecutively included in the RCT and treated with single-blastocyst transfer in mNC-FET without LPS until October 2022. Study participants were aged 18-41 years, had regular menstrual cycles (23–35 days) and underwent mNC-FET

treatment with the transfer of one good-quality (Gardner score ≥3BB), autologous blastocyst, vitrified on day 5 or 6 after oocyte retrieval. Oocyte recipients and couples undergoing preimplantation genetic testing were excluded. A full overview of inclusion and exclusion criteria has been published elsewhere (*Saupstad et al., 2019*).

FET protocol

FET was performed in mNC-FET. The timing of ovulation triggering was decided by ultrasonography. When the leading follicle reached 17 mm or more, human chorionic gonadotrophin (HCG; Ovitrelle, 250 μg) was administered at 22:00 hours the same evening. Transfer of one vitrified-warmed blastocyst was performed 6 or 7 days after ovulation triggering according to the randomization allocation. Serum HCG was measured 16 days after administration of the ovulation trigger and a pregnancy scan was performed during gestational week 8.

Progesterone measurements

On the day of blastocyst transfer, blood for testing was taken by venepuncture within 1 h before or after blastocyst transfer. All transfers were performed between 11:00 and 14:00 hours, 6 or 7 days after administration of the ovulation trigger. Blood sample analyses for serum progesterone concentrations were performed locally at each participating centre using immunoassays with an inter-assay variation of 10% or less. A full overview of the apparatus and assays used for progesterone analysis is depicted in Supplementary Table 1.

End-points

The objective of this study was to compare reproductive outcome rates between groups divided by progesterone concentrations of <29.0 and ≥29.0 nmol/l (25th percentile) on the day of blastocyst transfer. The primary end-point was the CPR, defined as a viable pregnancy beyond gestational week 7 (*Zegers-Hochschild et al., 2017*). Secondary outcomes were the pregnancy rate, defined as a serum HCG concentration ≥5 IU/l 16 days after the administration of the ovulation trigger, and pregnancy loss rate (PLR), defined as pregnancy loss detected before, or at the time of, the pregnancy scan in gestational week 8 (biochemical and clinical pregnancy loss).

Statistics

Statistical analyses were performed in IBM SPSS (IBM, USA). For the baseline and reproductive characteristics, continuous variables were expressed as medians and ranges, and categorical variables were expressed as frequencies and percentages. Progesterone concentrations on the day of blastocyst transfer were expressed as means and divided into percentiles. Site-specific progesterone percentiles were also calculated. To compare baseline and reproductive characteristics between women with progesterone concentrations <25th and ≥25th percentile, chi-squared tests for categorical variables and independent sample t-tests and Mann–Whitney *U*-tests for continuous variables were applied. Comparisons of reproductive outcomes below and at or above predefined progesterone cut-off concentrations (i.e. the 25th and 10th percentiles) were performed using chi-squared tests. To account for progesterone measurements at different time points, the dataset was split according to the day of blastocyst transfer (day 5 or day 6), after which comparisons of reproductive outcomes using chi-squared tests were repeated.

To adjust for potential confounding factors, multivariate binary logistic regression analyses were performed with clinical pregnancy as the dependent variable. Two multivariate regression models were tested including progesterone as a continuous and as a categorical (<25th and ≥25th percentile) covariate, respectively. The other covariates included in the regression analyses were: day of blastocyst transfer (6 or 7 days following the ovulation trigger), study site (1–8), male infertility (yes/no), female infertility (yes/no), single/female partner (yes/no), unknown reason for infertility (yes/no), body mass index (BMI; continuous), age (continuous), previous live birth (yes/no) and day of blastocyst vitrification (5 or 6). *P*-values <0.05 were considered statistically significant.

Ethics approval and trial registration

The original RCT was approved by the Danish Committee on Health Research Ethics (H-18025839, 6 September 2018), the Danish Medicines Agency (63569, 8 August 2018) and the Capital Region of Denmark, Research and Innovation, Legal Department (P-2019-670, 6 November 2019), and registered in EudraCT (2018-002207-34) and on ClinicalTrials.gov (NCT03795220).

RESULTS

Background characteristics

TABLE 1 demonstrates the background and reproductive characteristics among women with progesterone concentrations <25th or ≥25th percentile (<29.0 or ≥29.0 nmol/l). Women with progesterone concentrations below 29.0 nmol/l had a significantly higher BMI ($P = 0.001$) than women with progesterone concentrations ≥29.0 nmol/l. Significantly more women with progesterone concentrations below 29.0 nmol/l underwent blastocyst transfer and progesterone measurement at 6, compared with 7, days after the ovulation trigger ($P < 0.001$).

The proportion of single women/women with a female partner was significantly higher in the under 29.0 nmol/l group. These women had a significantly higher mean BMI than patients with other reasons for infertility (26.2 versus 23.6 kg/m^2; $P = 0.004$). In a logistic regression analysis with progesterone <29.0 versus ≥29.0 nmol/l as the dependent variable, and BMI and single/female partner as predictors, only BMI was significantly associated with progesterone concentration ($P = 0.002$). No other differences between the groups were observed.

Progesterone concentrations

Among the 209 women undergoing mNC-FET without LPS, the range of progesterone concentrations on the day of blastocyst transfer was 4.9–91.8 nmol/l (excluding one extreme outlier of 182.8 nmol/l). The 25th percentile was 29.0 nmol/l, and the 10th percentile was 22.5 nmol/l (TABLE 2). Site-specific progesterone ranges and percentiles can be found in Supplementary Tables 1 and 2. Mean (SD) progesterone concentrations on the day of blastocyst transfer were similar in women who did and did not achieve clinical pregnancy (38.5 [14.0] nmol/l versus 36.8 [12.4] nmol/l; $P = 0.350$). Comparable mean (SD) progesterone concentrations among women who did and did not achieve a clinical pregnancy were also seen when stratifying data according to blastocyst transfer on day 6 (34.2 [11.7] nmol/l versus 33.5 [11.9], $P = 0.758$); day 7 (42.3 [15.0] nmol/l versus 39.9 [12.1] nmol/l, $P = 0.360$).

Reproductive outcomes

The overall CPR was 36.8%. Reproductive outcomes after stratifying data according

TABLE 1 BACKGROUND AND REPRODUCTIVE CHARACTERISTICS IN WOMEN WITH SERUM PROGESTERONE CONCENTRATIONS <29.0 AND ≥29.0 NMOL/L ON THE DAY OF BLASTOCYST TRANSFER

Characteristics	Progesterone (nmol/l)		*P*-value
	<29.0 nmol/l	≥29.0 nmol/l	
Number	51	158	–
Age (years), mean (SD)	32.7 (3.6)	33.4 (4.0)	0.275
BMI (kg/m^2), median (IQR)	24.8 (22.1–28.5)	22.3 (20.5–25.7)	0.001
Cycle length (days), mean (SD)	28.8 (2.1)	28.4 (2.1)	0.245
AMH (pmol/l), median (IQR)	22.0 (12.5–34.5)	18.0 (12.0–27.0)	0.088
Duration of infertility (months), median (IQR)[a]	32 (24–41)	32 (24–36.5)	0.928
Cause of infertility, *n* (%)[b]			
Male factor[c]	22 (44)	82 (51.9)	0.330
Female factor[c]	7 (14)	35 (22.2)	0.211
Single/female partner [c]	10 (20.0)	11 (7.0)	0.008
Unknown[d]	16 (32.7)	45 (28.5)	0.576
Previous live birth, *n* (%)			
No	37 (72.5)	122 (77.2)	0.747
Yes	14 (27.5)	36 (22.8)	
Number of previous oocyte retrievals, median (range)	1 (1–3)	1 (1–3)	0.446
Number of previous FET cycles, median (range)	1 (1–≥4)	1 (1–≥4)	0.151
Method of fertilization *n*, (%)			
IVF	26 (51.0)	68 (43.0)	0.377
ICSI	25 (49.0)	87 (55.1)	
Days of culturing before vitrification, *n* (%)			
5 days	39 (76.5)	119 (75.3)	0.867
6 days	12 (23.5)	39 (24.7)	
Day of blastocyst transfer, *n* (%)[e]			
Day 6 after HCG trigger	37 (72.5)	63 (39.9)	<0.001
Day 7 after HCG trigger	14 (27.5)	95 (60.1)	

[a] *n* = 119, calculated for heterosexual couples only.
[b] *n* = 23 had multiple causes of infertility; therefore, the sum of the percentages within each group <29.0 and ≥29.0 nmol/l is >100.
[c] Denominator = 50.
[d] Denominator = 49.
[e] Corresponding to the day of serum progesterone measurement.
AMH, anti-Müllerian hormone; BMI, body mass index; FET, frozen embryo transfer; ICSI, intracytoplasmic sperm injection; IQR, interquartile range.

to the 25th or 10th serum progesterone percentile on the day of blastocyst transfer are shown in Figures 1 and 2, respectively. There were no statistically significant differences in reproductive outcome between the groups. The CPR for women with progesterone concentrations <29.0 or ≥29.0 nmol/l (25th percentile) was 35.3% and 37.3% ($P = 0.792$), respectively. The pregancy rate was 51.0% versus 54.8% ($P = 0.637$) and the PLR was 15.7% versus 17.2% ($P = 0.802$). For progesterone concentrations <22.5 or ≥22.5 nmol/l (10th percentile), the CPR was 36.8% versus 36.8% ($P = 1.000$), the pregancy rate was 47.4% versus 54.5% ($P = 0.552$) and the PLR was 10.5% versus 17.5% ($P = 0.441$). The results were similar when stratifying according to the day of blastocyst transfer, i.e. the day of progesterone measurement, 6 or 7 days after ovulation triggering (Supplementary Table 3, data shown for the 25th percentile). Site-specific CPRs in women grouped by progesterone concentration <29.0 and ≥29.0 nmol/l are found in Supplementary Table 1.

Adjusted multivariate logistic regression analyses demonstrated no association between progesterone concentrations on the day of blastocyst transfer and clinical pregnancy, when adding progesterone either as a continuous or a categorical (<29.0 versus ≥29.0 nmol/l) variable (TABLE 3, Supplementary Table 4). Effect estimates of progesterone concentrations were largely unaffected by adjustment.

TABLE 2 SERUM PROGESTERONE CONCENTRATIONS ON THE DAY OF BLASTOCYST TRANSFER IN MODIFIED NATURAL CYCLE FROZEN EMBRYO TRANSFER WITHOUT THE USE OF LUTEAL PHASE PROGESTERONE SUPPLEMENTATION

Parameter	Percentiles						
	5th	10th	25th	50th	75th	90th	95th
Progesterone (nmol/l)	19.2	22.5	29.0	36.7	43.8	54.0	60.6

DISCUSSION

This study is the first to investigate whether there is an association between serum progesterone concentrations on the day of single-blastocyst transfer and CPR in mNC-FET without LPS. It was found that the CPR was similar in women with progesterone concentrations <25th and ≥25th (29.0 nmol/l) or <10th and ≥10th (22.5 nmol/l) percentiles. Likewise, no significant differences in pregancy rate and PLR in women with progesterone concentrations <25th and 10th or ≥25th and 10th percentiles were found. The findings were confirmed in repeated analyses after stratifying data according to the day of blastocyst transfer, as well as in logistic regression analyses with or without adjustment for potential confounding factors.

In recent years, several studies, including one meta-analysis (*Labarta et al., 2021*; *Melo et al., 2021*), have come to the conclusion that serum progesterone concentrations below approximately 8–10 ng/ml (approximately 30 nmol/l) around the day of blastocyst transfer are associated with a reduced CPR or LBR in HRT-FET cycles. The aim of this study was

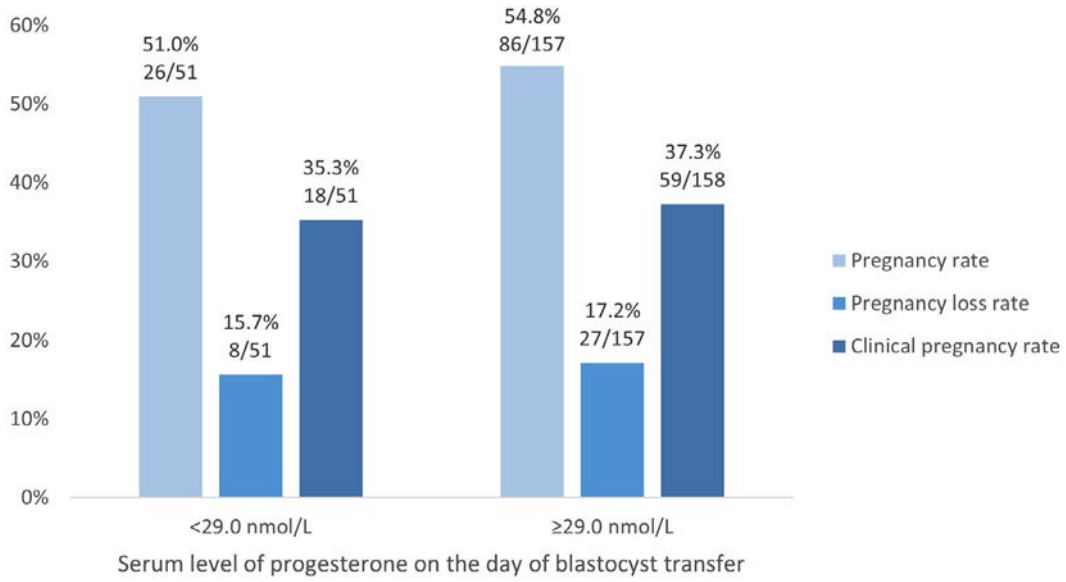


FIGURE 1 Comparison of reproductive outcomes according to serum progesterone concentrations <29.0 or ≥29.0 nmol/l. Data are shown as % (*n*/*N*). The cut-off value of 29.0 nmol/l corresponds to the 25th percentile in the present dataset.

to investigate whether the same principle applies to mNC-FET cycles without the administration of exogenous progesterone. Its results indicate that premises derived from studies on HRT-FET and tNC-FET cycles may not be applicable to mNC-FET treatment.

Reviewing the current literature, only two studies were found that had investigated the association between luteal phase progesterone concentrations and reproductive outcome in NC-FET without progesterone LPS (*Gaggiotti-Marre et al., 2020*; *Ramezanali et al., 2019*). Ramezanali and colleagues (*n* = 101) found no association between progesterone concentrations and CPR in mNC-FET cycles with cleavage-stage (day 2 or 3) embryo transfer, whereas Gaggiotti-Marre and co-workers (*n* = 294) reported a significantly lower CPR and LBR in women with progesterone concentrations below 10 ng/ml (31.8 mmol/l) on the day before blastocyst transfer in tNC-FET cycles.

In mNC-FET, ovulation is induced by the administration of an ovulation trigger, most commonly HCG. HCG acts through LH receptors expressed on luteinized theca and granulosa cells (*Bildik et al., 2020*) and has a relatively long half-life compared with LH. It can be postulated that the continuous, prolonged stimulation of the early corpus luteum by exogenous HCG serves as a mild form of LPS that, together with the physiological LH peak, secures sufficient corpus luteum activity. This may explain why, at this point, no association has been found between progesterone concentrations and reproductive outcomes in mNC-FET.

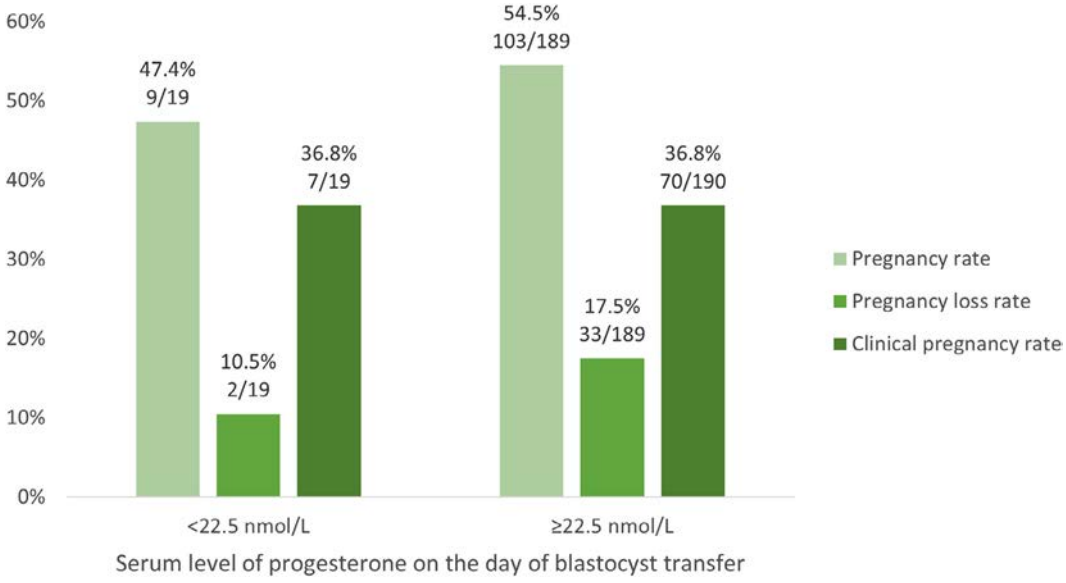


FIGURE 2 Comparison of reproductive outcomes according to serum progesterone concentrations <22.5 or ≥22.5 nmol/l. Data are shown as % (*n*/*N*). The cut-off value of 22.5 nmol/l corresponds to the 10th percentile in the present dataset.

TABLE 3 LOGISTIC REGRESSION ANALYSES EVALUATING THE IMPACT OF SERUM PROGESTERONE ON BLASTOCYST TRANSFER DAY ON THE CLINICAL PREGNANCY RATE IN MODIFIED NATURAL CYCLE FROZEN EMBRYO TRANSFER WITHOUT EXOGENOUS PROGESTERONE SUPPLEMENTATION

Parameter	aOR (95% CI)[a]
Serum progesterone concentrations <29.0 or ≥29.0 nmol/l, categorical[b]	0.87 (0.39–1.93)
Serum progesterone concentrations, continuous	1.01 (0.99–1.04)

[a] Adjusted odds ratio. Adjusted for age, body mass index, site, cause of infertility, previous birth, day of vitrification and day of blastocyst transfer.

[b] Using <29.0 nmol/l as the reference value.

LPS is widely used in both tNC-FET and mNC-FET. Two recent systematic reviews and meta-analyses have investigated the effect of LPS on NC-FET outcome (*Mizrachi et al., 2021*; *Seol et al., 2020*). Seol and collaborators found no association between the use of LPS and CPR, while Mizrachi and colleagues reported an increase in CPR with the use of LPS. Out of the seven studies included in the reviews, two small RCT (Eftakhar et al., *n* = 102; Horowitz et al., *n* = 59) were performed in mNC-FET cycles and reported no effect of LPS on reproductive outcomes (*Eftakhar et al., 2013*; *Horowitz et al., 2021*).

In contrast, a fresh from the press RCT evaluating the effect of progesterone supplementation on pregnancy outcome in tNC-FET found that supplementation with exogenous progesterone from the day of embryo transfer improved LBR (*Wånggren et al., 2022*). However, when assessing the relationship between progesterone concentrations and pregnancy among women not receiving progesterone supplementation, the authors found no significant correlation between LBR and progesterone serum concentration on the day of embryo transfer. The authors argue that a possible positive effect of exogenous progesterone may not be reflected in a single serum progesterone measurement. In summary, whether women treated using NC-FET benefit from the use of LPS remains controversial and the effect of LPS may differ in tNC-FET and mNC-FET cycles.

Among the strengths of the current study is the fact that the patient population is derived from an RCT. The authors acknowledge the fact that data derived from a sub-study of an RCT do not have the same power to provide unbiased estimates of effect as the RCT itself. However, they argue that data prospectively collected from an RCT, where participants are randomized to treatment rather than treated according to clinician or patient preferences, do not have the same risk of selection bias as is seen in prospective or retrospective cohort studies. All women allocated to mNC-FET without LPS until the day of data extraction were included in the sub-study.

One of the limitations of this study is that the dataset might be underpowered to detect an effect of progesterone concentration at the day of blastocyst transfer on pregnancy outcomes, increasing the risk of a type II error. Nevertheless, the CPR rates were similar in the groups with progesterone <29.0 or ≥29.0 nmol/l, making a type II error less likely. Comparing this study with a similar study on tNC-FET by Gaggiotti-Mare and colleagues (*n* = 294; *Gaggiotti-Marre et al., 2020*), the sample size of the current study is modest (*n* = 209), and it cannot be ruled out that the results might have been different with a larger number of participants. In particular, the multivariate analysis including progesterone concentrations <25th and ≥25th percentiles (dichotomous) might be subject to overadjustment as the sample size for the <25th percentile is small. To overcome this problem, a multivariate analysis including progesterone as a continuous variable, including the full sample size, was performed, showing an almost unchanged odds ratio after adjustment (Supplementary Table 4).

While women in the <29.0 and ≥29.0 nmol/l progesterone groups were similar on many parameters, the BMI significantly differed between the two groups as women in the under 29.0 nmol/l group were more likely to have a higher mean BMI. Lower concentrations of progesterone around the day of blastocyst transfer in overweight women in NC-FET with (*Brady et al., 2014*) and without (*Gaggiotti-Marre et al., 2020*; *Mumusoglu et al., 2023*) LPS have previously been reported and may be explained by an increased distribution volume.

The day of blastocyst transfer, and thus the day of progesterone measurement, differed significantly between women with progesterone concentrations <29.0 and ≥29.0 nmol/l, with a higher rate of day 7 transfer in the ≥29.0 nmol/l progesterone group. The natural rise in progesterone concentration after ovulation may explain this difference. To account for progesterone measurements at two different time-points, the study population was stratified according to the day of blastocyst transfer, 6 or 7 days after ovulation triggering. Between-group analyses were repeated after stratification, and the results did not change significantly (Supplementary Table 3). In addition, associations between progesterone concentrations and CPR remained insignificant in adjusted analyses including the day of blastocyst transfer as a covariate.

While the day of progesterone measurement varied, the timing of blood drawing was fixed within 1 h before or after transfer, i.e. between 11:00 and 14:00 hours. A standardized time of measurement is a strength, as progesterone concentrations tend to fluctuate throughout the day (*Filicori et al., 1984*; *Thomsen et al., 2018*). However, single measurements of serum progesterone concentration may not be representative of progesterone concentrations throughout the day, and a time span of 3 h for blood drawing may be too wide.

For logistical reasons, progesterone analyses were performed locally at each trial site. Using different immunoassays for progesterone analysis may compromise the comparability of progesterone concentrations between sites, and centralized blood sample analysis would have been preferable. However, as shown in Supplementary Tables 1 and 2, progesterone concentrations were comparable between sites. No differences were found in CPR between the <29.0 and ≥29.0 nmol/l progesterone groups at any study site; however, the multivariate regression analysis indicated that the CPR might differ between study sites. Both analyses are, however, subject to small sample sizes at several study sites and should be interpreted with caution.

While an absence of evidence does not necessarily mean evidence of an absence of effect, it was not possible to identify a significant association between progesterone concentrations on the day of blastocyst transfer and CPR, pregancy rate or PLR in mNC-FET without LPS. The authors recognize the limitations of the current study, including a modest sample size and the use of different assays for progesterone analysis. Therefore, much is yet to be uncovered regarding the impact of progesterone concentrations, as well as the use of LPS, in NC-FET and particularly mNC-FET cycles. The data emphasize that results from HRT-FET and tNC-FET cycles might not be applicable to mNC-FET, stressing the importance of distinguishing between different endometrial preparation protocols prior to FET in research as well as clinical practice.

DATA AVAILABILITY

Data will be made available on request.

ACKNOWLEDGEMENTS

The authors would like to thank chemist Lærke Andersen from the Department of Clinical Biochemistry at Aalborg University Hospital for help clarifying the different progesterone assays used at the different trial sites.

FUNDING

The study was fully supported by Rigshospitalet's Research Foundation and an unrestricted grant from Gedeon Richter (Forward Grant, FORWARD2018_5).

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103862.

REFERENCES

Alsbjerg, B., Thomsen, L., Elbaek, H.O., Laursen, R., Povlsen, B.B., Haahr, T., Humaidan, P., 2020. Can combining vaginal and rectal progesterone achieve the optimum progesterone range required for implantation in the HRT-FET model? Reprod. Biomed. Online 40, 805–811. https://doi.org/10.1016/j.rbmo.2020.02.007.

Basnayake, S.K., Volovsky, M., Rombauts, L., Osianlis, T., Vollenhoven, B., Healey, M., 2018. Progesterone concentrations and dosage with frozen embryo transfers – What's best? Aust. N. Z. J. Obstet. Gynaecol. 58, 533–538. https://doi.org/10.1111/ajo.12757.

Bildik, G., Akin, N., Esmaeilian, Y., Hela, F., Yakin, K., Onder, T., Urman, B., Oktem, O., 2020. hCG Improves Luteal Function and Promotes Progesterone Output through the Activation of JNK Pathway in the Luteal Granulosa Cells of the Stimulated IVF Cycles†. Biol. Reprod. 102, 1270–1280. https://doi.org/10.1093/biolre/ioaa034.

Brady, P.C., Kaser, D.J., Ginsburg, E.S., Ashby, R.K., Missmer, S.A., Correia, K.F., Racowsky, C., 2014. Serum progesterone concentration on day of embryo transfer in donor oocyte cycles. J. Assist. Reprod. Genet. 31, 569–575. https://doi.org/10.1007/s10815-014-0199-y.

Busnelli, A., Schirripa, I., Fedele, F., Bulfoni, A., Levi-Setti, P.E., 2022. Obstetric and perinatal outcomes following programmed compared to natural frozen-thawed embryo transfer cycles: a systematic review and meta-analysis. Hum. Reprod. 37, 1619–1641. https://doi.org/10.1093/huPLRep/deac073.

Cédrin-Durnerin, I., Isnard, T., Mahdjoub, S., Sonigo, C., Seroka, A., Comtet, M., Herbemont, C., Sifer, C., Grynberg, M., 2019. Serum progesterone concentration and live birth rate in frozen–thawed embryo transfers with hormonally prepared endometrium. Reprod. Biomed. Online 38, 472–480. https://doi.org/10.1016/j.rbmo.2018.11.026.

Eftekhar, M., Rahsepar, M., Rahmani, E., 2013. Effect of progesterone supplementation on natural frozen-thawed embryo transfer cycles: A randomized controlled trial. Int. J. Fertil. Steril. 7, 13–20.

Filicori, M., Butler, J.P., Crowley, W.F., 1984. Neuroendocrine regulation of the corpus luteum in the human. Evidence for pulsatile progesterone secretion. J. Clin. Invest. 73, 1638–1647. https://doi.org/10.1172/JCI111370.

Gaggiotti-Marre, S., Alvarez, M., González-Foruria, I., Parriego, M., Garcia, S., Martínez, F., Barri, P.N., Polyzos, N.P., Coroleu, B., 2020. Low progesterone levels on the day before natural cycle frozen embryo transfer are negatively associated with live birth rates. Hum. Reprod. 35, 1623–1629. https://doi.org/10.1093/huPLRep/deaa092.

Ghobara, T., Gelbaya, T.A., Ayeleke, R.O., 2017. Cycle regimens for frozen-thawed embryo transfer. Cochrane database Syst. Rev. 7, CD003414. https://doi.org/10.1002/14651858.CD003414.pub3.

Groenewoud, E.R., Cantineau, A.E.P., Kollen, B.J., Macklon, N.S., Cohlen, B.J., 2017. What is the optimal means of preparing the endometrium in frozen-thawed embryo transfer cycles? A systematic review and meta-analysis. Hum. Reprod. Update 22, 255–261. https://doi.org/10.1093/humupd/dmt030.

Horowitz, E., Mizrachi, Y., Finkelstein, M., Farhi, J., Shalev, A., Gold, E., Raziel, A., Weissman, A., 2021. A randomized controlled trial of vaginal progesterone for luteal phase support in modified natural cycle–frozen embryo transfer. Gynecol. Endocrinol. 37, 792–797. https://doi.org/10.1080/09513590.2020.1854717.

Labarta, E., Mariani, G., Holtmann, N., Celada, P., Remohí, J., Bosch, E., 2017. Low serum progesterone on the day of embryo transfer is associated with a diminished ongoing pregnancy rate in oocyte donation cycles after artificial endometrial preparation: A prospective study. Hum. Reprod. 32, 2437–2442. https://doi.org/10.1093/huPLRep/dex316.

Labarta, E., Mariani, G., Paolelli, S., Rodriguez-Varela, C., Vidal, C., Giles, J., Bellver, J., Cruz, F., Marzal, A., Celada, P., Olmo, I., Alamá, P., Remohi, J., Bosch, E., 2021. Impact of low serum progesterone levels on the day of embryo transfer on pregnancy outcome: A prospective cohort study in artificial cycles with vaginal progesterone. Hum. Reprod. 36, 683–692. https://doi.org/10.1093/huPLRep/deaa322.

Melo, P., Chung, Y., Pickering, O., Price, M.J., Fishel, S., Khairy, M., Kingsland, C., Lowe, P., Petsas, G., Rajkhowa, M., Sephton, V., Tozer, A., Wood, S., Labarta, E., Wilcox, M., Devall, A., Gallos, I., Coomarasamy, A., 2021. Serum luteal phase progesterone in women undergoing frozen embryo transfer in assisted conception: a systematic review and meta-analysis. Fertil. Steril. 116, 1534–1556. https://doi.org/10.1016/j.fertnstert.2021.07.002.

Mizrachi, Y., Horowitz, E., Ganer Herman, H., Farhi, J., Raziel, A., Weissman, A., 2021. Should women receive luteal support following natural cycle frozen embryo transfer? A systematic review and meta-analysis. Hum. Reprod. Update 27, 643–650. https://doi.org/10.1093/humupd/dmab011.

Moreno-Sepulveda, J., Espinós, J.J., Checa, M.A., 2021. Lower risk of adverse perinatal outcomes in natural versus artificial frozen-thawed embryo transfer cycles: a systematic review and meta-analysis. Reprod. Biomed. Online 42, 1131–1145. https://doi.org/10.1016/j.rbmo.2021.03.002.

Mumusoglu, S., Erden, M., Ozbek, I.Y., Ince, O., Esteves, S.C., Humaidan, P., Yarali, H., 2023. The true natural cycle frozen embryo transfer – impact of patient and follicular phase characteristics on serum progesterone levels one day prior to warmed blastocyst transfer. Reprod. Biol. Endocrinol. 21, 86. https://doi.org/10.1186/s12958-023-01136-z.

Ramezanali, F., Arabipoor, A., Hafezi, M., Salman-Yazdi, R., Zolfaghari, Z., Asharfi, M., 2019. Serum estradiol level on trigger day impacts clinical pregnancy rate in modified natural frozen embryo transfer cycles. Int. J. Gynecol. Obstet. 145, 312–318. https://doi.org/10.1002/ijgo.12806.

Rienzi, L., Gracia, C., Maggiulli, R., LaBarbera, A.R., Kaser, D.J., Ubaldi, F.M., Vanderpoel, S., Racowsky, C., 2017. Oocyte, embryo and blastocyst cryopreservation in ART: systematic review and meta-analysis comparing slow-freezing versus vitrification to produce evidence for the development of global guidance. Hum. Reprod. Update 23, 139–155. https://doi.org/10.1093/humupd/dmw038.

Saupstad, M., Freiesleben, N.L.C., Skouby, S.O., Andersen, L.F., Knudsen, U.B., Petersen, K.B., Husth, M., Egeberg, A., Petersen, M.R., Ziebe, S., Andersen, A.N., Løssl, K., Pinborg, A., 2019. Preparation of the endometrium and timing of blastocyst transfer in modified natural cycle frozen-thawed embryo transfers (mNC-FET): A

study protocol for a randomised controlled multicentre trial. BMJ Open 9, 1–7. https://doi.org/10.1136/bmjopen-2019-031811.

Seol, A., Shim, Y.J., Kim, S.W., Kim, S.K., Lee, J.R., Jee, B.C., Suh, C.S., Kim, S.H., 2020. Effect of luteal phase support with vaginal progesterone on pregnancy outcomes in natural frozen embryo transfer cycles: A meta-analysis. Clin. Exp. Reprod. Med. 47, 147–152. https://doi.org/10.5653/CERM.2019.03132.

Stormlund, S., Sopa, N., Zedeler, A., Bogstad, J., Prætorius, L., Nielsen, H.S., Kitlinski, M.L., Skouby, S.O., Mikkelsen, A.L., Spangmose, A.L., Jeppesen, J.V., Khatibi, A., la Cour Freiesleben, N., Ziebe, S., Polyzos, N.P., Bergh, C., Humaidan, P., Andersen, A.N., Løssl, K., Pinborg, A., 2020. Freeze-all versus fresh blastocyst transfer strategy during in vitro fertilisation in women with regular menstrual cycles: multicentre randomised controlled trial. BMJ 370, m2519. https://doi.org/10.1136/bmj.m2519.

Thomsen, L.H., Kesmodel, U.S., Andersen, C.Y., Humaidan, P., 2018. Daytime Variation in Serum Progesterone During the Mid-Luteal Phase in Women Undergoing In Vitro Fertilization Treatment. Front. Endocrinol. (Lausanne). 9, 92. https://doi.org/10.3389/fendo.2018.00092.

Volovsky, M., Pakes, C., Rozen, G., Polyakov, A., 2020. Do serum progesterone levels on day of embryo transfer influence pregnancy outcomes in artificial frozen-thaw cycles? J. Assist. Reprod. Genet. 37, 1129–1135. https://doi.org/10.1007/s10815-020-01713-w.

von Versen-Höynck, F., Narasimhan, P., Selamet Tierney, E.S., Martinez, N., Conrad, K.P., Baker, V.L., Winn, V.D., 2019. Absent or Excessive Corpus Luteum Number Is Associated With Altered Maternal Vascular Health in Early Pregnancy. Hypertension 73, 680–690. https://doi.org/10.1161/HYPERTENSIONAHA.118.12046.

Wånggren, K., Dahlgren Granbom, M., Iliadis, S.I., Gudmundsson, J., Stavreus-Evers, A., 2022. Progesterone supplementation in natural cycles improves live birth rates after embryo transfer of frozen-thawed embryos-a randomized controlled trial. Hum. Reprod. 37, 2366–2374. https://doi.org/10.1093/huPLRep/deac181.

Wu, H., Zhou, P., Lin, X., Wang, S., Zhang, S., 2021. Endometrial preparation for frozen—thawed embryo transfer cycles: a systematic review and network meta-analysis. J. Assist. Reprod. Genet. 38, 1913–1926. https://doi.org/10.1007/s10815-021-02125-0.

Yovich, J.L., Conceicao, J.L., Stanger, J.D., Hinchliffe, P.M., Keane, K.N., 2015. Mid-luteal serum progesterone concentrations govern implantation rates for cryopreserved embryo transfers conducted under hormone replacement. Reprod. Biomed. Online 31, 180–191. https://doi.org/10.1016/j.rbmo.2015.05.005.

Zaat, T., Zagers, M., Mol, F., Goddijn, M., van Wely, M., Mastenbroek, S., 2021. Fresh versus frozen embryo transfers in assisted reproduction. Cochrane Database Syst. Rev. 2, CD011184. https://doi.org/10.1002/14651858.CD011184.pub3.

Zegers-Hochschild, F., Adamson, G.D., Dyer, S., Racowsky, C., de Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., Cooke, I.D., Simpson, J.L., van der Poel, S., 2017. The International Glossary on Infertility and Fertility Care, 2017. Fertil. Steril. 108, 393–406. https://doi.org/10.1016/j.fertnstert.2017.06.005.

Received 18 July 2023; received in revised form 29 January 2024; accepted 30 January 2024.

RBMO

ELSEVIER

REVIEW

What reproductive follow-up for adolescent and young women after cancer? A review

BIOGRAPHY

Christine Decanter is Head of the Assisted Reproductive Technology and Fertility Preservation Centre of Lille University Hospital (France). She has set up an oncofertility observatory, and her areas of research are oncofertility, anti-Müllerian hormone, oocyte quality and polycystic ovary syndrome.

C. Decanter[a,b,*], E. Elefant[c,d], C. Poirot[e,f], B. Courbiere[g,h]

KEY MESSAGE

Fertility after cancer is highly challenging for young adult cancer survivors. Dedicated follow-up of ovarian function in adolescent and young adult women is crucial to evaluate chemotherapy-induced ovarian toxicity, to refine fertility preservation strategies, and to help in fertility counselling and future family planning.

ABSTRACT

Fertility capacity has been shown to be one of the main concerns of young cancer survivors. Gonadotoxic treatments may lead to both premature ovarian failure and/or infertility. This review aimed to define which, and when, reproductive indicators should be followed-up to help doctors to counsel patients regarding their fertility and ovarian function, and to determine if a second stage of fertility preservation after the end of cancer treatment is clinically relevant. Longitudinal assessment of anti-Müllerian hormone (AMH) concentrations during cancer treatment indicates the degree of follicular depletion, and allows discrimination between low and high gonadotoxic treatments. Sustained low AMH concentrations after treatment, especially in the case of alkylating protocols, may reduce the duration of the conception window significantly, and expose the patient to the risk of premature ovarian failure. It remains unknown whether this may impact further fertility capacity because of the lack of systematic follow-up of adolescent and young adult (AYA) women after chemo-radiotherapy. It appears that dedicated reproductive follow-up of AYA women under cancer treatment is needed to refine fertility preservation strategies, and to determine if low AMH concentrations after treatment impact the chance of pregnancy in this specific survivor population.

INTRODUCTION

Diagnostic and therapeutic progress in the oncologic area has made it possible to achieve a survival rate of 70–80% in children, adolescents and young adults (AYA) (*Tonorezos et al., 2022*). Despite increasing overall survival rates, therapy-associated side effects continue to be a major concern. Many centres have established specific AYA cancer programmes to take into consideration the specificities of this subgroup in terms of information, education and global care. Publications have defined the age range of AYA as 15–25, 15–35 or 15–38 to 39 years old. Fertility after cancer is obviously a crucial factor of the young survivor's quality of life, and counselling in fertility/fertility preservation techniques is now an essential part of the oncologic care programme (*Lewin et al., 2017*; *Tonorezos et al., 2022*).

[a] CHU Lille, ART and Fertility Preservation Department, Lille, France
[b] ONCOLille Cancer Institute, Lille, France
[c] Reference Centre for Teratogenic Agents, Hospital Armand Trousseau, Paris, France
[d] Faculty of Medicine, Sorbonne University, Paris, France
[e] Department of Haematology, Adolescents and Young Adults, Fertility Preservation, Assistance Publique des Hôpitaux de Paris, Saint Louis Hospital, Paris, France
[f] Medecine Sorbonne University, Paris, France
[g] Department of Gynaecology–Obstetric and Reproductive Medicine–Fertility Preservation, AP-HM, Hôpital La Conception, Marseille, France
[h] Aix-Marseille Université, Avignon Université, Avignon, France

KEY WORDS
Cancer
Fertility
Fertility preservation
AMH
Chemotherapy
AYA

*Corresponding author. E-mail address: christine.decanter@chu-lille.fr (C. Decanter). https://doi.org/10.1016/j.rbmo.2024.103891 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

Longitudinal studies of ovarian function through menstrual cycle profiles and anti-Müllerian hormone (AMH) concentration dynamics in AYA women have played a determinant role in better comprehension of the mechanisms of chemotherapy-induced ovarian toxicity, allowing chemotherapeutic protocols to be classified according to their degree of ovarian toxicity. It is now well established that cycle recovery alone after treatment is insufficient to assess protocol toxicity and predict future fertility. AMH is secreted by granulosa cells of pre-antral and antral follicles, and is recognized to be the most specific, accurate, reproducible and non-invasive marker of ovarian follicular content (*Dewailly et al., 2014*). It has also been shown to be a real-time indicator of follicular depletion and renewal in patients treated with chemotherapy (*Peigné et al., 2014*). Longitudinal studies of AMH variations in young women undergoing chemotherapy make it possible to develop decision trees for fertility preservation, and to provide best advice for future family planning. This review aimed to define which, and when, reproductive indicators should be followed-up to help doctors to counsel patients regarding their fertility and ovarian function, and to determine if a second stage of fertility preservation at an appropriate time after cancer treatment is clinically relevant.

MONITORING OVARIAN FUNCTION DURING AND AFTER TREATMENT

AMH trajectories according to protocol and age

Lymphoma and breast cancer are the most represented in longitudinal studies, and are a good model for the influence of age on ovarian function as lymphoma affects patients in their 20s, whereas breast cancer affects patients in their 30s.

In women with breast cancer or lymphoma, longitudinal studies highlighted a follicular depletion phase, reflected by a rapid and profound decrease in AMH concentration, as early as the first cycle of chemotherapy, regardless of protocol. This was followed, after the end of treatment, by slow, partial recovery of AMH concentration in the case of an alkylating regimen (*Anderson et al., 2017*, *2018b*; *Decanter et al., 2010*, *2014*, *2018*, *2021*; *Demeestere et al., 2021*; *Dezellus et al., 2017*; *Dillon et al., 2013*; *Mailliez et al., 2022*; *Perdrix et al., 2017*) or by rapid, complete renewal in the case of a non-alkylating regime (*Anderson et al., 2018a*,*b*; *Decanter et al., 2010*, *2021*).

Patients with lymphoma treated with the adriamycin-bleomycin-vinblastine - dacarbazine protocol recovered, ad integrum, their initial AMH concentrations by 6-month post-treatment follow-up (*Anderson et al., 2018b*; *Decanter et al., 2010*, *2021*; *Demesteere et al., 2021*). Conversely, AMH concentrations in women treated with an alkylating regimen never returned to pretreatment values, and remained low or undetectable throughout the follow-up period, suggesting deep, extensive ovarian follicular injury (*Anderson et al., 2018a*,b; *Decanter et al., 2010*; *2021*; *Demesteere et al., 2021*).

A recent study compared longitudinal variations in AMH concentrations in AYA patients (age 15–24 years) with non-AYA patients (age 25–35 years), and demonstrated that young age does not protect against the effects of alkylating protocols on ovarian reserve; as such, it is essential to propose fertility preservation (*Decanter et al., 2021*). Indeed, AMH remained undetectable in 20% of women aged 15–24 years, including 5% with premature ovarian insufficiency (POI) at 2-year follow-up after the end of treatment, versus 45% and 10% in women aged 25–35 years.

In patients with breast cancer, despite a lower cumulative dose of cyclophosphamide, sequential anthracycline and taxane-based chemotherapy also induced a large decrease in the AMH concentration in women aged <35 years, and this appears to reinforce ovarian toxicity in comparison with an anthracycline alkylating regimen (*Dezellus et al., 2017*; *Lambertini et al., 2019*; *Mailliez et al., 2022*; *Perdrix et al., 2017*). A recent study of 125 patients with breast cancer aged ≤35 years found that 35% had ultra-low (≤5 pmol/l) or even undetectable AMH concentrations 12 and 24 months after the end of treatment, and 5% had POI (*Mailliez et al., 2022*). In all studies, the extent of recovery of AMH concentration was marginal, and none of the patients demonstrated a return to their pretreatment values. Patients with undetectable post-treatment AMH concentrations were at higher risk of definitive amenorrhoea (*Anderson et al., 2017*), although several studies reported patients who did recover spontaneous menstrual function despite low or undetectable AMH concentrations (*Decanter et al., 2018*; *Dezellus et al., 2017*; *Lambertini et al., 2019*, *Mailliez et al., 2022*; *Perdrix et al., 2017*).

Pretreatment AMH concentration and age were strongly associated with the degree of ovarian recovery in lymphoma and breast cancer survivors (*Anderson et al., 2013*; *Decanter et al., 2021*).

Timing and long-term evolution of ovarian function recovery

In most cases, regardless of the type of cancer, AMH concentrations increase within the first 2 years of follow-up, and menses resume. Thereafter, AMH concentrations plateau at significantly lower values than pretreatment (*Cameron et al., 2019*; *Decanter et al., 2021*; *Perdrix et al., 2017*; *Su et al., 2020*) and compared with controls of the same age, especially in patients for whom the cyclophosphamide equivalent dose exceeded 6000 mg/m^2 (*Drechsel et al., 2023a*). The impact of the cyclophosphamide equivalent dose on fertility impairment was confirmed recently in a large European case–control study involving child and AYA cancer survivors (PancareLIFE), with a cut-off of 7000 mg/m^2 (*van den Berg et al., 2021*). Limited data on menstrual function and AMH concentrations after 5 years of follow-up are available. One longitudinal, cross-sectional study investigated ovarian function in 763 women treated at 18–39 years of age for up to 15 years after cancer diagnosis. Patients who received low/moderate gonadotoxic protocols exhibited a long-lasting plateau in AMH concentration for 10–15 years before declining. In contrast, for patients who received highly gonadotoxic protocols, AMH concentrations declined shortly after peaking at 2–3 years, suggesting a higher risk of POI or infertility (*Su et al., 2020*).

Despite a high incidence of extremely low or undetectable AMH concentrations in patients who received an alkylating regimen, most AYA survivors return to spontaneous menses within 1–2 years of follow-up. However, the bleeding pattern in AYA survivors, especially in those aged 15–25 years, is difficult to interpret, and AMH concentrations fluctuate significantly and physiologically during adolescence and up to 24 years of age (*Kelsey et al.2012*). In most studies, the return of menses was not systematically related to AMH concentration, at least when measured using a conventional assay (*Decanter et al.,*

2021; Dezellus et al., 2017; Mailliez et al., 2022; Su et al., 2020). Conversely, use of a pico-AMH assay, which is 10-fold more sensitive, made it possible to discriminate between young adult patients who will recover spontaneous menses from those who will not among the breast cancer population (Chai et al., 2014; Decanter et al., 2014, 2018). It is of particular importance to consider interassay variability, especially at the lower limit of detection, rending it difficult to determine cut-off values to predict POI. Likewise, in patients with breast cancer in their 30s, post-treatment AMH concentrations were not related to the short-term occurrence of pregnancy (Hamy et al., 2016), as in the general population where low AMH concentrations are not predictive of time to pregnancy (Hagen et al., 2012; Steiner et al., 2017). Data regarding associations between indicators of ovarian reserve and pregnancy occurrence are still lacking in survivors treated at 15–25 years of age, but it is likely that very low AMH concentrations may reduce the conception window, and affect the reproductive lifespan in younger patients, especially as mechanisms other than acute follicle apoptosis during chemotherapy may occur, such as accelerated activation of primordial follicles (Kalich-Philosoph et al., 2013; Roness et al., 2014), and ovarian stroma and vasculature injury causing fibrosis (Meirow et al., 2007), both of which induce a second phase of follicular loss, and the latter may impact long-term evolution of ovarian function (Shai et al., 2022; Spears et al., 2019). Interestingly, the only study that compared the rate of decline in AMH concentration in young cancer survivors, at a median time from cancer therapy of 9.9 years, with age-matched controls did not find any significant differences, except for lower AMH concentrations throughout follow-up in the survivors (Cameron et al., 2019). In other studies, AMH concentration has been shown to predict POI in patients treated for cancer (Anderson et al., 2022; Thomas-Teinturier et al., 2015; Vathaire et al., 2015).

FERTILITY IN AYA CANCER SURVIVORS

All data in the literature regarding female fertility after cancer are from retrospective studies. The main limitation in interpreting the results is the lack of knowledge of the exact number of patients who tried to conceive compared with the numbers of pregnancies and deliveries. There are also other biases, notably patient fear of relapsing if they become pregnant, especially in cases with a history of breast cancer. Therefore, rates of infertility described in the literature are possibly overstated.

Various cancers

A retrospective epidemiological study reported a high incidence of infertility (40% at 35 years of age) and POI (5–10% depending on cancer type) in 1041 women aged 18–40 years who had been treated for haematological malignancies, breast cancer or gastrointestinal tumours (Letourneau et al., 2012). Similar data were observed in survivors who had been treated for various cancers in childhood or adolescence (Gerstl et al., 2018a). Another large retrospective study from the Scottish registry involved patients aged ≤39 years at diagnosis of various cancers (n = 23,201, including 10,271 nulliparous women), and matched for age with patients from the general population (Anderson et al., 2018b). Cancer survivors achieved fewer pregnancies with a standardized incidence rate of 0.62 (95% CI 0.60–0.63) for all cancer types, with larger decreases for breast, cervical and brain cancers, and leukaemia.

Haematological malignancies

A recent narrative review of 75 articles undertaken to study reproductive ability in survivors of Hodgkin lymphoma reported that the incidence of POI was 6–34% (median 9%), depending on age at treatment and protocol received (Drechsel et al., 2023a,b). A recent meta-analysis of 14 studies, including 744 young women who received autogenic or allogeneic bone marrow transplants for haematological malignancies, highlighted a lower conception rate compared with the general population (25% and 22%, respectively), with a miscarriage rate of approximately 10% (Gerstl et al., 2019). Likewise, data from the Scottish registry reported lower conception rates for patients (aged < 39 years) diagnosed with Hodgkin lymphoma, non-Hodgkin lymphoma and leukaemia following cancer treatment compared with age-matched controls (Anderson et al., 2018a). A recent study conducted in survivors of Hodgkin lymphoma, treated between 2000 and 2015, highlighted parenthood rates similar to those of the general population but with higher usage of assisted reproductive technology for the first pregnancy (Ovlisen et al., 2021). The same results have been shown in a population of non-Hodgkin lymphoma survivors, except for those with the more aggressive clinical form, treated with the highest dose of alkylating agents, where the childbirth rate was lower than the comparators (hazard ratio 0.43, 95% CI 0.48–1.04) during the first 3 years post diagnosis (Entrop et al., 2023). Of note is a greater decrease, in recent years, in the chance of pregnancy in lymphoma survivors due to more aggressive treatments, such as BEACOPP, to reduce the incidence of relapse (Anderson et al., 2018; Behringer et al., 2005; Entrop et al., 2023; Ovlisen et al., 2021). Conversely, recent strategies of therapeutic 'de-escalation' in the case of a satisfactory response to positron emission scanning show more favourable chances of pregnancy (Demeestere et al., 2021).

Breast cancer

A systematic review followed by a meta-analysis in young women with breast cancer treated with chemotherapy reported a 14% lower overall pregnancy rate in subgroup analysis of case–control or matched cohort studies (n = 1287 patients) (Gerstl et al., 2018b). The incidence of pregnancy after breast cancer is 40% lower than in the general population. A large study (27,556 survivors including 1240 women; age of treatment between 16 and 45 years) reported a 67% lower chance of pregnancy in patients treated for breast cancer, compared with the general population, after adjusting for age, socio-educational level, parity and time of diagnosis (Stensheim et al., 2011). Chemotherapeutic regimens containing anthracyclines were associated with a greater probability of pregnancy compared with taxane-containing regimens (hazard ratio 4.75, 95% CI 1.76–12.8; P = 0.002), with the latter being the most widely used (Hamy et al., 2016). A few studies have reported pregnancies in some patients despite low or undetectable AMH concentrations after cancer treatment (Anderson et al., 2018, Dezellus et al., 2017; Hamy et al., 2016; Mailliez et al., 2022). This is consistent with the lack of predictive value of AMH concentration for pregnancy in healthy women (Steiner et al., 2017), but data in the specific group of cancer survivors are too scarce to make formal conclusions.

WHAT TYPE OF REPRODUCTIVE CARE IN AYA SURVIVORS?

Which biomarker to follow and when?

FIGURE 1 shows the suggested follow-up of reproductive indicators in AYA survivors who received an alkylating regimen.

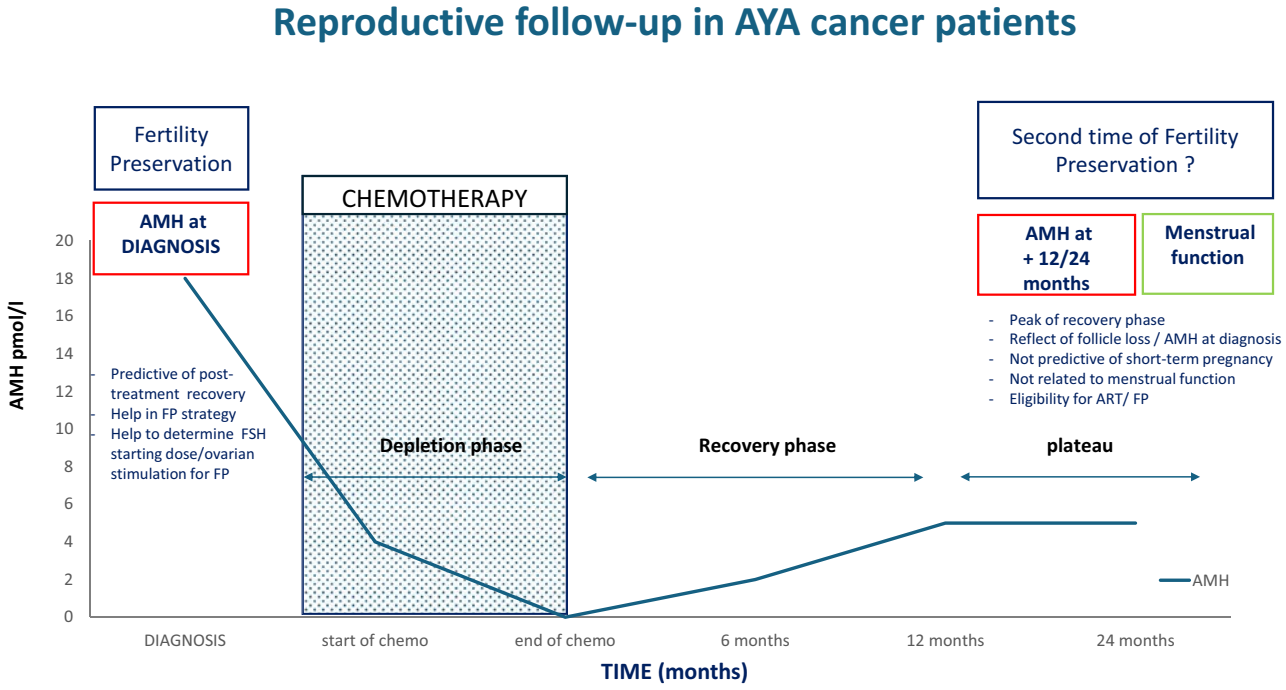


FIGURE 1 Suggested follow-up of reproductive indicators in adolescent and young adult cancer survivors who received an alkylating chemotherapy regimen. Anti-Müllerian hormone (AMH) is the best indicator of follicular depletion and renewal during and after chemotherapy. Menstrual function should be checked at each follow-up visit. FP, fertility preservation; ART, assisted reproductive technology.

Studies tracking AMH concentrations before and after chemotherapy in young women treated for breast cancer or lymphoma allowed assessment of the degree of follicular loss and potential for recovery based on the protocol received, patient's age and pretreatment AMH concentration. Although there are several arguments in favour of premature ovarian senescence due to chemo-radiotherapy, it is still unclear whether and how AMH concentrations after cancer contribute to defining the reproductive aging stage of AYA. In addition, in the youngest patients, aged 15–25 years, bleeding pattern and fluctuations of ovarian reserve indicators must be interpreted with caution. Nevertheless, measuring AMH concentrations at the peak of folliculogenesis recovery (12–24 months after the end of treatment) and thereafter once every 1 or 2 years may help doctors in fertility/fertility preservation counselling, and patients for future family planning. Furthermore, post-treatment AMH concentrations may help in choosing the gonadotrophin starting dose for controlled ovarian stimulation if ART or fertility preservation is required. Recent French recommendations from the National Cancer Institute and the Biomedicine Agency advise systematic follow-up by serial AMH measurements before commencing chemotherapy, at the end of treatment, and 12 and/or 24 months later, as well as menstrual function tracking (*Rives et al., 2022*). This dedicated long-term follow-up is crucial, especially in the youngest patients for whom the delay between their cancer treatment and pregnancy will be the longest. Tracking reproductive indicators longitudinally will also provide new insights into the efficiency and safety of fertility preservation procedures performed at cancer diagnosis, which are needed for future perspectives on the elaboration of evidence-based fertility preservation strategies.

Should we propose a second time of fertility preservation at a distance of cancer treatment?

Despite the undeniable progress and development of fertility preservation techniques, the number of patients eligible before the start of cancer treatment remains low. Logistical difficulties, alteration of the general state, contraindication to stimulation treatments or laparoscopy, urgency of treatment, or lack of dedicated information may explain the low level of recruitment. Similarly, even when the procedures have been applied, it is not common to benefit, in a single collection, from enough oocytes to offer consistent chances of subsequent pregnancy and live birth. Cryopreservation of oocytes by vitrification is a recent technique which has proven effectiveness and safety, at least in healthy women aged <35 years. Regarding patients with cancer, we are at the very beginning of return for reutilization due to POI or infertility. To date, the largest study of reutilization of oocytes cryopreserved in oncologic patients aged <35 years before cancer treatment ($n = 42$) showed a reutilization rate of approximately 8%, and cumulative live birth rates of 35% and 43% if at least eight or 10 mature oocytes were cryopreserved (*Cobo et al., 2021*). Likewise, a recent meta-analysis reported live birth rates after reutilizing cryopreserved embryos or oocytes of 41% and 32%, respectively (*Fraison et al., 2023*). Considering reutilization of cryopreserved ovarian tissue, the spontaneous live birth rate was 33%, and the live birth rate if IVF was required was 21% (*Fraison et al., 2023*). All these arguments, together with the significant reduction in the conception window, the frequent postponement of the pregnancy plan from 2–5 years after the end of treatment in certain types of cancer, and the risk of further POI, may lead to the proposal of a second attempt at fertility preservation, by oocyte accumulation, after agreement with the oncologist. The concept of oocyte accumulation has been proposed in conventional ART and in fertility

preservation programmes in women with predicted low ovarian response in relation to low AMH concentrations, with the aim of optimizing the oocyte yield and allowing further chances of pregnancy if reutilization is needed (*Cobo et al., 2012*; *Legrand et al., 2021*). In patients with cancer, these procedures should be performed at a suitable time after treatment (i.e. 1–2 years later) to benefit from the peak of ovarian follicular recovery, and to avoid the period of maximum incidence of disease relapse. One to three collection cycles can be proposed to obtain an optimal number of oocytes (i.e. approximately 10). The prospect of a second stage of fertility preservation after treatment should not lead to neglecting the proposal of fertility preservation before treatment because, unfortunately, a significant proportion of patients will not recover sufficient ovarian reserve to be eligible for fertility preservation, or will have POI after treatment.

CONCLUSION

This literature review determined the procedure for monitoring reproductive indicators in AYA cancer survivors. AMH concentration appears to be the best real-time indicator of follicular depletion and renewal during and after cancer treatment over 2–3 years of follow-up. There is still a crucial need for long-term follow-up to be able to elaborate formal recommendations regarding further fertility preservation strategies and future family planning. This is of particular interest in the youngest patients (age 15–25 years), for whom the reproductive lifespan is likely to be the most affected.

DATA AVAILABILITY

No data was used for the research described in the article.

ACKNOWLEDGEMENTS

The authors wish to thank Andy Thomson, Thomson Language Consultants, for proofreading.

AUTHOR CONTRIBUTIONS

CD and BC performed the review of the literature and selection of relevant articles. CD and BC drafted the manuscript. CD wrote the final manuscript. CP and EE checked the selected articles and revised the manuscript for intellectual content. All authors approved the final version of the manuscript.

REFERENCES

Anderson, RA, Rosendahl, M, Kelsey, TW, Cameron, DA., 2013. Pretreatment anti-Müllerian hormone predicts for loss of ovarian function after chemotherapy for early breast cancer. Eur J Cancer 49 (16), 3404–3411. https://doi.org/10.1016/j.ejca.2013.07.014 Nov.

Anderson, RA, Mansi, J, Coleman, RE, Adamson, DJ, Leonard, RC., 2017. The utility of anti-Müllerian hormone in the diagnosis and prediction of loss of ovarian function following chemotherapy for early breast cancer. Eur J Cancer 87, 58–64. https://doi.org/10.1016/j.ejca.2017.10.001.

Anderson, RA, Brewster, DH, Wood, R, Nowell, S, Fischbacher, C, Kelsey, TW, Wallace, WHB., 2018a. The impact of cancer on subsequent chance of pregnancy: a population-based analysis. Hum Reprod 33 (7), 1281–1290. https://doi.org/10.1093/humrep/dey216 Jul 1.

Anderson, R.A., Remedios, R., Kirkwood, A.A., Patrick, P., Stevens, L., Clifton-Hadley, L., Roberts, T., Hatton, C., Kalakonda, N., Milligan, D.W., McKay, P., Rowntree, C., Scott, F.M., Johnson, P.W.M., 2018b. Determinants of ovarian function after response-adapted therapy in patients with advanced Hodgkin's lymphoma (RATHL): a secondary analysis of a randomised phase 3 trial. Lancet Oncol 19 (10), 1328–1337. https://doi.org/10.1016/S1470-2045(18)30500-X.

Anderson, RA, Cameron, D, Clatot, F, Demeestere, I, Lambertini, M, Nelson, SM, Peccatori, F., 2022. Anti-Müllerian hormone as a marker of ovarian reserve and premature ovarian insufficiency in children and women with cancer: a systematic review. Hum Reprod Update 28 (3), 417–434. https://doi.org/10.1093/humupd/dmac004 May 2.

Behringer, K, Mueller, H, Goergen, H, Thielen, I, Eibl, AD, Stumpf, V, Wessels, C, Wiehlpütz, M, Rosenbrock, J, Halbsguth, T, Reiners, KS, Schober, T, Renno, JH, von Wolff, M, van der Ven, K, Kuehr, M, Fuchs, M, Diehl, V, Engert, A, Borchmann, P, 2013. Gonadal function and fertility in survivors after Hodgkin lymphoma treatment within the German Hodgkin Study Group HD13 to HD15 trials. J Clin Oncol 31 (2), 231–239. https://doi.org/10.1200/JCO.2012.44.3721 Epub 2012 Nov 13. PMID: 23150709.

Cameron, K, Sammel, MD, Prewitt, M, Gracia, C., 2019. Differential Rates of Change in Measures of Ovarian Reserve in Young Cancer Survivors Across the Reproductive Lifespan. J Clin Endocrinol Metab 104 (5), 1813–1822. https://doi.org/10.1210/jc.2018-02257 May 1.

Chai, J, Howie, AF, Cameron, DA, Anderson, RA., 2014. A highly-sensitive anti-Müllerian hormone assay improves analysis of ovarian function following chemotherapy for early breast cancer. Eur J Cancer 50 (14), 2367–2374. https://doi.org/10.1016/j.ejca.2014.06.011 Sep.

Cobo, A, García-Velasco, JA, Remohí, J, Pellicer, A., 2021. Oocyte vitrification for fertility preservation for both medical and nonmedical reasons. Fertil Steril 115 (5), 1091–1101. https://doi.org/10.1016/j.fertnstert.2021.02.006 May.

Cobo, A, Garrido, N, Crespo, J, José, R, Pellicer, A., 2012. Accumulation of oocytes: a new strategy for managing low-responder patients. Reprod Biomed Online 24 (4), 424–432. https://doi.org/10.1016/j.rbmo.2011.12.012 Apr.

Decanter, C., Morschhauser, F., Pigny, P., Lefebvre, C., Gallo, C., Dewailly, D., 2010. Anti-Mullerian hormone follow-up in young women treated by chemotherapy for lymphoma: preliminary results. Reprod Biomed Online 20 (2), 280–285. https://doi.org/10.1016/j.rbmo.2009.11.010.

Decanter, C., Peigne, M., Mailliez, A., Morschhauser, F., Dassonneville, A., Dewailly, D., Pigny, P., 2014. Toward a better follow-up of ovarian recovery in young women after chemotherapy with a hypersensitive antimullerian hormone assay. Fertil Steril 102 (2), 483–487. https://doi.org/10.1016/j.fertnstert.2014.05.014.

Decanter, C, Cloquet, M, Dassonneville, A, D'Orazio, E, Mailliez, A, Pigny, P, 2018. Different patterns of ovarian recovery after cancer treatment suggest various individual ovarian susceptibilities to chemotherapy. Reprod Biomed Online 36 (6), 711–718. https://doi.org/10.1016/j.rbmo.2018.02.004.

Decanter, C, Delepine, J, Behal, H, Manier, S, Bruno, B, Barbatti, M, Robin, C, Labreuche, J, Morschhauser, F, Pigny, P., 2021. Longitudinal study of AMH variations in 122 Adolescents and Young Adults (AYA) and non-AYA lymphoma patients to evaluate the chemo-induced ovarian toxicity to further personalise fertility preservation counselling. Hum Reprod 36 (10), 2743–2752. https://doi.org/10.1093/humrep/deab189 Sep 18.

Demeestere, I, Racape, J, Dechene, J, Dupuis, J, Morschhauser, F, De Wilde, V, Lazarovici, J, Ghesquieres, H, Touati, M, Sibon, D, Alexis, M, Gac, AC, Moatti, H, Virelizier, E, Maisonneuve, H, Pranger, D, Houot, R, Fornecker, LM, Tempescul, A, André, M, Casasnovas, RO, 2021. Gonadal Function Recovery in Patients With Advanced Hodgkin Lymphoma Treated With a PET-Adapted Regimen: Prospective Analysis of a Randomized Phase III Trial (AHL2011). J Clin Oncol 39 (29), 3251–3260. https://doi.org/10.1200/JCO.21.00068 Oct 10.

Dewailly, D., Andersen, C.Y., Balen, A., Broekmans, F., Dilaver, N., Fanchin, R., Griesinger, G., Kelsey, T.W., La Marca, A., Lambalk, C., Mason, H., Nelson, S.M., Visser, J.A., Wallace, W.H., Anderson, R.A., 2014. The physiology and clinical utility of anti-Mullerian hormone in women. Hum Reprod Update 20 (3), 370–385. https://doi.org/10.1093/humupd/dmt062.

Dezellus, A, Barriere, P, Campone, M, Lemanski, C, Vanlemmens, L, Mignot, L, Delozier, T, Levy, C, Bendavid, C, Debled, M, Bachelot, T, Jouannaud, C, Loustalot, C, Mouret-Reynier, MA, Gallais-Umbert, A, Masson, D, Freour, T, 2017. Prospective evaluation of serum anti-Müllerian hormone dynamics in 250 women of reproductive age treated with chemotherapy for breast cancer. Eur J Cancer 79, 72–80. https://doi.org/10.1016/j.ejca.2017.03.035.

Dillon, K.E., Sammel, M.D., Prewitt, M., Ginsberg, J.P., Walker, D., Mersereau, J.E., Gosiengfiao, Y., Gracia, C.R., 2013. Pretreatment antimullerian hormone levels determine rate of posttherapy ovarian reserve recovery: acute changes in ovarian reserve during and after chemotherapy. Fertil Steril 99 (2), 477–483. https://doi.org/10.1016/j.fertnstert.2012.09.039.

Drechsel, KCE, Broer, SL, Stoutjesdijk, FS, Twisk, JWR, van den Berg, MH, Lambalk, CB, van Leeuwen, FE, Overbeek, A, van den Heuvel-Eibrink, MM, van Dorp, W, de Vries, ACH, Loonen, JJ, van der Pal, HJ, Kremer, LC, Tissing, WJ, Versluys, B, Kaspers, GJL, van Dulmen-den Broeder, E, Veening, MA, LATER-VEVO study group, 2023a. Clinical and self-reported markers of reproductive function in female survivors of childhood Hodgkin lymphoma. J Cancer Res Clin Oncol 149 (15), 13677–13695. https://doi.org/10.1007/s00432-023-05035-z.

Drechsel, KCE, Pilon, MCF, Stoutjesdijk, F, Meivis, S, Schoonmade, LJ, Wallace, WHB, van Dulmen-den Broeder, E, Beishuizen, A, Kaspers, GJL, Broer, SL, Veening, MA., 2023b. Reproductive ability in survivors of childhood, adolescent, and young adult Hodgkin lymphoma: a review. Hum Reprod Update 29 (4), 486–517. https://doi.org/10.1093/humupd/dmad002 Jul 5.

Entrop, JP, Weibull, CE, Smedby, KE, Jakobsen, LH, Øvlisen, AK, Glimelius, I, Marklund, A, Larsen, TS, Holte, H, Fosså, A, Smeland, KB, El-Galaly, TC, Eloranta, S., 2023. Reproduction patterns among non-Hodgkin lymphoma survivors by subtype in Sweden, Denmark and Norway: A population-based matched cohort study. Br J Haematol. https://doi.org/10.1111/bjh.18938 Jun 16Epub ahead of print. PMID: 37325886.

Fraison, E, Huberlant, S, Labrune, E, Cavalieri, M, Montagut, M, Brugnon, F, Courbiere, B., 2023. Live birth rate after female fertility preservation for cancer or haematopoietic stem cell transplantation: a systematic review and meta-analysis of the three main techniques; embryo, oocyte and ovarian tissue cryopreservation. Hum Reprod 38 (3), 489–502. https://doi.org/10.1093/humrep/deac249 Mar 1.

Gerstl, B, Sullivan, E, Chong, S, Chia, D, Wand, H, 2018a. Anazodo A Reproductive Outcomes After a Childhood and Adolescent Young Adult Cancer Diagnosis in Female Cancer Survivors: A Systematic Review and Meta-analysis. J Adolesc Young Adult Oncol. https://doi.org/10.1089/jayao.2018.0036 Nov 16.

Gerstl, B, Sullivan, E, Ives, A, Saunders, C, Wand, H, Anazodo, A., 2018b. Pregnancy Outcomes After a Breast Cancer Diagnosis: A Systematic Review and Meta-analysis. Clin Breast Cancer 18 (1), e79–e88. https://doi.org/10.1016/j.clbc.2017.06.016 Feb.

Gerstl BSullivan, E, Koch, J, Wand, H, Ives, A, Mitchell, R, Hamad, N, Anazodo, A, 2019. Reproductive outcomes following a stem cell transplant for a haematological malignancy in female cancer survivors: a systematic review and meta-analysis. Support Care Cancer 27 (12), 4451–4460. https://doi.org/10.1371/journal.pone.0256497 Dec.

Hagen, C.P., Vestergaard, S., Juul, A., Skakkebaek, N.E., Andersson, A.M., Main, K.M., Hjollund, N.H., Ernst, E., Bonde, J.P., Anderson, R.A., Jensen, T.K., 2012. Low concentration of circulating antimullerian hormone is not predictive of reduced fecundability in young healthy women: a prospective cohort study. Fertil Steril 98 (6). https://doi.org/10.1016/j.fertnstert.2012.08.008 1602-1608 e1602.

Hamy, AS, Porcher, R, Eskenazi, S, Cuvier, C, Giacchetti, S, Coussy, F, Hocini, H, Tournant, B, Perret, F, Bonfils, S, Charvériat, P, Lacorte, JM, Espie, M., 2016. Anti-Müllerian hormone in breast cancer patients treated with chemotherapy: a retrospective evaluation of subsequent pregnancies. Reprod Biomed Online 32 (3), 299–307. https://doi.org/10.1016/j.rbmo.2015.12.008 Mar.

Kalich-Philosoph, L, Roness, H, Carmely, A, Fishel-Bartal, M, Ligumsky, H, Paglin, S, Wolf, I, Kanety, H, Sredni, B, Meirow, D., 2013. Cyclophosphamide triggers follicle activation and "burnout"; AS101 prevents follicle loss and preserves fertility. Sci Transl Med 5 (185), 185ra62. https://doi.org/10.1126/scitranslmed.3005402 May 15.

Kelsey, TW, Anderson, RA, Wright, P, Nelson, SM, Wallace, WH., 2012. Data-driven assessment of the human ovarian reserve. Mol Hum Reprod 18 (2), 79–87. https://doi.org/10.1093/molehr/gar059 Feb.

Lambertini, M, Olympios, N, Lequesne, J, Calbrix, C, Fontanilles, M, Loeb, A, Leheurteur, M, Demeestere, I, Di Fiore, F, Perdrix, A, Clatot, F., 2019. Impact of Taxanes, Endocrine Therapy, and Deleterious Germline BRCA Mutations on Anti-müllerian Hormone Levels in Early Breast Cancer Patients Treated With Anthracycline- and Cyclophosphamide-Based Chemotherapy. Front Oncol 9, 575. https://doi.org/10.3389/fonc.2019.00575 Jul 12.

Legrand, C, Keller, L, Collinet, P, Barbotin, AL, Béhal, H, Rubod, C, Decanter, C., 2021. Oocyte accumulation for fertility preservation in women with benign ovarian tumours with a history of previous surgery, multiple or large cysts. Reprod Biomed Online 43 (2), 205–214. https://doi.org/10.1016/j.rbmo.2021.04.020 Aug.

Letourneau, JM, Ebbel, EE, Katz, PP, et al., 2012. Acute ovarian failure underestimates age-specific reproductive impairment for young women undergoing chemotherapy for cancer. Cancer 118 (7), 1933–1939. https://doi.org/10.1002/cncr.26403.

Lewin, J, Ma, JMZ, Mitchell, L, Tam, S, Puri, N, Stephens, D, Srikanthan, A, Bedard, P, Razak, A, Crump, M, Warr, D, Giuliani, M, Gupta, A., 2017. The positive effect of a dedicated adolescent and young adult fertility program on the rates of documentation of therapy-associated infertility risk and fertility preservation options. Support Care Cancer 25 (6), 1915–1922. https://doi.org/10.1007/s00520-017-3597-8 Jun.

Mailliez, A, Pigny, P, Bogart, E, Keller, L, D'Orazio, E, Vanseymortier, M, Le Deley, MC, Decanter, C, 2022. Is ovarian recovery after chemotherapy in young patients with early breast cancer influenced by controlled ovarian hyperstimulation for fertility preservation or tumor characteristics? Results of a prospective study in 126 patients. Int J Cancer 150 (11), 1850–1860. https://doi.org/10.1016/j.fertnstert.2014.05.014 Jun 1.

Meirow, D, Dor, J, Kaufman, B, Shrim, A, Rabinovici, J, Schiff, E, Raanani, H, Levron, J, Fridman, E., 2007. Cortical fibrosis and blood-vessels damage in human ovaries exposed to chemotherapy. Potential mechanisms of ovarian injury. Hum Reprod 22 (6), 1626–1633. https://doi.org/10.1093/humrep/dem027 Jun.

Øvlisen, AK, Jakobsen, LH, Eloranta, S, Kragholm, KH, Hutchings, M, Frederiksen, H, Kamper, P, Dahl-Sørensen, RB, Stoltenberg, D, Weibull, CE, Entrop, JP, Glimelius, I, Smedby, KE, Torp-Pedersen, C, Severinsen, MT, El-Galaly, TC., 2021. Parenthood Rates and Use of Assisted Reproductive Techniques in Younger Hodgkin Lymphoma Survivors: A Danish Population-Based Study. J Clin Oncol 39 (31), 3463–3472. https://doi.org/10.1200/JCO.21.00357 Nov 1.

Peigne, M., Decanter, C, 2014. Serum AMH level as a marker of acute and long-term effects of chemotherapy on the ovarian follicular content: a systematic review. Reprod Biol Endocrinol 12, 26. https://doi.org/10.1186/1477-7827-12-26.

Perdrix, A, Saint-Ghislain, M, Degremont, M, et al., 2017. Influence of adjuvant chemotherapy on anti-Müllerian hormone in women below 35 years treated for early breast cancer. Reprod Biomed Online 35 (4), 468–474. https://doi.org/10.1016/j.rbmo.2017.06.005.

Rives, N, Courbière, B, Almont, T, Kassab, D, Berger, C, Grynberg, M, Papaxanthos, A, Decanter, C, Elefant, E, Dhedin, N, Barraud-Lange, V, Béranger, MC, Demoor-Goldschmidt, C, Frédérique, N, Bergère, M, Gabrel, L, Duperray, M, Vermel, C, Hoog-Labouret, N, Pibarot, M, Provansal, M, Quéro, L, Lejeune, H, Methorst, C, Saias, J, Véronique-Baudin, J, Giscard d'Estaing, S, Farsi, F, Poirot, C, Huyghe, É., 2022. What should be done in terms of fertility preservation for patients with cancer? The French 2021 guidelines. Eur J Cancer 173, 146–166. https://doi.org/10.1016/j.ejca.2022.05.013 Sep.

Roness, H, Kalich-Philosoph, L, Meirow, D., 2014. Prevention of chemotherapy-induced ovarian damage: possible roles for hormonal and non-hormonal attenuating agents. Hum Reprod Update 20 (5), 759–774. https://doi.org/10.1093/humupd/dmu019 Sep-Oct.

Shai, D, Aviel-Ronen, S, Spector, I, Raanani, H, Shapira, M, Gat, I, Roness, H, Meirow, D., 2022. Ovaries of patients recently treated with alkylating agent chemotherapy indicate the presence of acute follicle activation, elucidating its role among other proposed mechanisms of follicle loss. Fertil Steril 115 (5), 1239–1249. https://doi.org/10.1016/j.fertnstert.2020.11.040 MayEpub 2021 Jan 20. PMID: 33485607.

Spears, N, Lopes, F, Stefansdottir, A, Rossi, V, De Felici, M, Anderson, RA, Klinger, FG., 2019. Ovarian damage from chemotherapy and current approaches to its protection. Hum Reprod Update 25 (6), 673–693. https://doi.org/10.1093/humupd/dmz027 Nov 5.

Steiner, A.Z., Pritchard, D., Stanczyk, F.Z., Kesner, J.S., Meadows, J.W., Herring, A.H., Baird, D.D., 2017. Association Between Biomarkers of Ovarian Reserve and Infertility Among Older Women of Reproductive Age. JAMA 318 (14), 1367–1376. https://doi.org/10.1001/jama.2017.14588.

Stensheim, H, Cvancarova, M, Møller, B, Fosså, SD., 2011. Pregnancy after adolescent and adult cancer: a population-based matched cohort study. Int J Cancer 129 (5), 1225–1236. https://doi.org/10.1002/ijc.26045 Sep 1.

Su, HI, Kwan, B, Whitcomb, BW, Shliakhsitsava, K, Dietz, AC, Stark, SS, Martinez, E, Sluss, PM, Sammel, MD, Natarajan, L., 2020. Modeling Variation in the Reproductive Lifespan of Female Adolescent and Young Adult Cancer Survivors Using AMH. J Clin Endocrinol Metab 105 (8), 2740–2751. https://doi.org/10.1210/clinem/dgaa172 Aug 1.

Thomas-Teinturier, C, Allodji, RS, Svetlova, E, Frey, MA, Oberlin, O, Millischer, AE, Epelboin, S, Decanter, C, Pacquement, H, Tabone, MD, Sudour-Bonnange, H, Baruchel, A, Lahlou, N, De Vathaire, F, 2015. Ovarian reserve after treatment with alkylating agents during childhood. Hum Reprod 30 (6), 1437–1446. https://doi.org/10.1093/humrep/dev060 Jun. Epub 2015 Mar 23. PMID: 25801499.

Tonorezos, ES, Cohn, RJ, Glaser, AW, Lewin, J, Poon, E, Wakefield, CE, Oeffinger, KC., 2022. Long-term care for people treated for cancer during childhood and adolescence. Lancet 399 (10334), 1561–1572. https://doi.org/10.1016/S0140-6736(22)00460-3 Apr 16. PMID: 35430023; PMCID: PMC9082556.

Vathaire, F., 2015. Ovarian reserve after treatment with alkylating agents during childhood. Hum Reprod 30 (6), 1437–1446. https://doi.org/10.1093/humrep/dev060 Jun.

van den Berg, MH, van Dijk, M, Byrne, J, Berger, C, Dirksen, U, Winther, JF, Fossa, SD, Grabow, D, Grandage, VL, Haupt, R, van den Heuvel-Eibrink, MM, Kaiser, M, Kepak, T, van der Kooi, ALF, Kremer, LCM, Kruseova, J, Lambalk, CB, van Leeuwen, FE, Leiper, A, Modan-Moses, D, Spix, C, Twisk, JWR, Ronckers, CM, Kaatsch, P, van Dulmen-den Broeder, E, PanCareLIFE Consortium, 2021. Treatment-related fertility impairment in long-term female childhood, adolescent and young adult cancer survivors: investigating dose-effect relationships in a European case-control study (PanCareLIFE). Hum Reprod 36 (6), 1561–1573. https://doi.org/10.1093/humrep/deab035 May 17.

Received 1 August 2023; received in revised form 2 January 2024; accepted 9 February 2024.

RBMO

ARTICLE

Gene associations of lipid traits, lipid-lowering drug-target genes and endometriosis

BIOGRAPHY

Ge Zhou, PhD, is Associate Professor, Chief Physician and Master's Degree Supervisor at the Affiliated Hospital of Nanjing University of Traditional Chinese Medicine (China). Her main research interests are assisted reproductive technology and diminished ovarian reserve.

Ge Zhou[1,†], Jin Ren[1,3,†], Qiuyan Huang[1,3], Xiaowei Nie[1], Xingli Tong[1], Ya wen Cui[1,3], Rongkui Hu[4,*], Qi Yao[2,*]

KEY MESSAGE

This study found a potential bidirectional causal association between endometriosis and dyslipidaemia. The development of lipid-lowering drugs to treat endometriosis may be of clinical interest.

ABSTRACT

Research question: Does the observed correlation between dyslipidaemia and endometriosis indicate a bidirectional causal association?

Design: Bidirectional Mendelian randomization was used to investigate the causal association between lipid traits and endometriosis. Drug-target Mendelian randomization was used to explore potential drug-target genes for managing endometriosis. In cases where lipid-mediated effects via specific drug targets were significant, aggregate analyses, such as summary-data-based Mendelian randomization and colocalization methods, were introduced to validate the outcomes. Mediation analyses supplemented these evaluations.

Results: The bidirectional Mendelian randomization results suggested that genetically predicted triglyceride (OR 1.15, 95% CI 1.08–1.23), high-density lipoprotein cholesterol (OR 0.87, 95% CI 0.81–0.94), low-density lipoprotein cholesterol (OR 1.20, 95% CI 1.06–1.34) and apolipoprotein A (OR 0.90, 95% CI 0.83–0.96) concentrations were causally associated with endometriosis. Reverse Mendelian randomization results revealed that genetically proxied endometriosis was causally associated with triglyceride concentration (OR 1.02, 95% CI 1.01–1.02). In drug-target Mendelian randomization, genetic mimicry in proprotein convertase subtilisin/kexin type 9 (*PCSK9*) (OR 1.40, 95% CI 1.13–1.72), apolipoprotein B (*APOB*) (OR 1.49, 95% CI 1.21–1.86) and angiopoietin-related protein 3 (*ANGPTL3*) (OR 1.57, 95% CI 1.14–2.16) was significantly associated with the risk of endometriosis stages 1–2.

Conclusion: There is a potential bidirectional causal association between endometriosis and dyslipidaemia. Genetic mimicry of *PCSK9*, *APOB* and *ANGPTL3* is associated with the risk of early-stage endometriosis. The development of lipid-lowering drugs to treat endometriosis is of potential clinical interest.

[1] Department of Reproductive Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Chinese Medicine, Nanjing, China
[2] Department of Pathology and Pathophysiology, Nanjing University of Chinese Medicine, Nanjing, Jiangsu Province, China
[3] The First College of Clinical Medical, Nanjing University of Chinese Medicine, Nanjing, China
[4] Gynaecology Department, Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Chinese Medicine, Nanjing, China
† These authors contributed equally.

KEY WORDS
Dyslipidaemia
Endometriosis
Target gene
Mendelian randomization

*Corresponding authors. E-mail addresses: fsyy00637@njucm.edu.cn (R. Hu), qiqiyao@126.com (Q. Yao). https://doi.org/10.1016/j.rbmo.2024.103856 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

INTRODUCTION

Endometriosis is the colonization of endometrial tissue outside the uterine cavity, which progresses with cyclical menstruation. Globally, it impacts approximately 8–10% of women in their childbearing years, and is a significant contributor to female infertility (*Bulun et al., 2019*; *Zondervan et al., 2020*). Endometriosis can develop insidiously as early as adolescence, and persists throughout the reproductive years. Clinical manifestations encompass chronic pelvic pain, dysmenorrhoea and deep dyspareunia (*Vercellini et al., 2014*). In the absence of effective pharmaceutical interventions, many women resort to recurrent use of analgesics (*Guo, 2023*). Given the extremity of surgical intervention in early-stage endometriosis and the multitude of side effects associated with long-term hormone therapy, early identification and modification of the risk factors for endometriosis have emerged as imperative strategies to mitigate its impact.

Prospective cohort studies have elucidated a positive epidemiological association between endometriosis and hypercholesterolaemia (*Mu et al., 2017*). However, the causality and direction of this association remain unconfirmed. Similarly, cross-sectional studies have revealed co-occurrence of endometriosis and dyslipidaemia (*Liang et al., 2022*), but the observation inevitably prompts enquiries into the directional aspect of this association. Moreover, the associations between various lipid profiles and endometriosis are contradictory. For instance, a positive association between triglycerides, low-density lipoprotein cholesterol (LDL-C) and endometriosis was reported in a cross-sectional study (*Melo et al., 2010*). Another prospective study found that the association between endometriosis and triglycerides was insignificant, whereas LDL-C exhibited a negative correlation with endometriosis (*Gibran et al., 2017*). These discrepancies may stem from inherent limitations in observational studies, including potential biases and confounding variables.

Despite the ambiguity in elucidating the association between endometriosis and dyslipidaemia, research into new treatments continues. Studies in baboons have demonstrated that 3-hydroxy-3-methylglutaryl CoA reductase (HMGCR) inhibitors, such as simvastatin, can reduce the risk of endometriosis by inhibiting the mevalonate pathway (*Taylor et al., 2017*). Recent studies have suggested a potential benefit of short-term atorvastatin treatment in women with endometriosis (*Dillon et al., 2022*). However, the absence of randomized controlled trials means that the repurposing of lipid-lowering drugs for the treatment of endometriosis remains theoretical.

Mendelian randomization, leveraging data from genome-wide association studies (GWAS), offers a novel approach in epidemiological research, circumventing issues such as residual confounding and reverse causation that are common in traditional studies. Given that genetic instruments are allocated randomly during meiosis (*Emdin et al., 2017*), Mendelian randomization can be conceptualized as a quasi-randomized natural experiment, seamlessly implemented within the same ancestry group. The bidirectional Mendelian randomization method serves to explore the causality between lipid traits and endometriosis. The summary-data-based Mendelian randomization (SMR) method uses pooled-level data from GWAS and expression quantitative trait loci (eQTL) studies to test for pleiotropic associations between gene expression levels and complex traits of interest, and the results are examined using heterogeneity in dependent instruments (HEIDI). Drug-target Mendelian randomization, an extension of univariable Mendelian randomization, identifies instrumental variables for eQTL or protein quantitative trait loci (pQTL) (*Huang et al., 2021*). This approach can mirror drug action mechanisms, providing insights into potential clinical applications. The present study used various Mendelian randomization techniques to investigate the causal relationship between lipid traits and endometriosis, and to assess the impact of lipid-lowering drug targets on the condition.

METHODS

This study was designed rigorously in alignment with the Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian randomization checklist (*Skrivankova et al., 2021*). Summary-level data from GWAS and eQTL studies were used, detailed in Supplementary Tables S1 and S2. All primary materials were approved by an institutional review board, and written informed consent was obtained from all participants. This paper is a secondary study and did not require further ethical review.

Data source

Single-nucleotide polymorphisms (SNP) significantly associated with triglycerides, high-density lipoprotein cholesterol (HDL-C), LDL-C, apolipoprotein A (APOA) and apolipoprotein B (APOB) were selected to delineate the various lipid traits (Supplementary Tables S3–S7). Genetic associations were derived from UK Biobank, a repository boasting the most recent and extensive compilation of summary-level GWAS data related to lipid traits. The baseline data from UK Biobank encompass nearly half a million participants recruited between 2006 and 2010, aged 40–69 years.

Case–control GWAS data for endometriosis were derived from FinnGen project R8, adhering to the International Classification of Diseases, 10th Revision. The FinnGen project provides granular data, categorizing endometriosis based on the revised American Society of Reproductive Medicine (rASRM) staging and lesion sites. For further analysis, the summary data on rASRM endometriosis stages 1–2 and stages 3–4 were stratified to analyse the causal association between lipid traits and different stages of endometriosis. Moreover, various endometriotic sites were selected to explore the causal association between lipid traits and the lesion location, specifically ovarian endometriosis and pelvic peritoneum endometriosis. Meta-analysis of GWAS data was adjusted for age and sex, ensuring accuracy. The datasets showed no participant overlap.

Most gene-encoding targets of lipid-lowering drugs were identified in the Drug Bank database, focusing on three categories: LDL-C-lowering drug targets [*HMGCR*, proprotein convertase subtilisin/kexin type 9 (*PCSK9*), Niemann-Pick C1-like protein 1 (*NPC1L1*), LDL receptor (*LDLR*) and apolipoprotein B (*APOB*)]; triglyceride-lowering drug target [angiopoietin-related protein 3 (*ANGPTL3*)]; and HDL-C-enhancing drug target [cholesteryl ester transfer protein (*CETP*)] (*Pahan, 2006*). SNP are shown in Supplementary Table S8.

Genetic instrument selection

To delineate exposure traits, the screening criteria for instrumental variables were GWAS *P*-values of <5e-08 and a clumping threshold of $r^2 < 0.001$ within a 1-Mb

genomic region (linkage disequilibrium) in univariable Mendelian randomization. Gene-target screening methods from the existing literature were used to proxy the exposure of drug targets, and SNP within 100-kb genomic region windows of each drug-target gene that were associated with lipid traits with a GWAS *P*-value of <5e-08 were identified. A further selection of SNP was allowed to weaken the linkage disequilibrium (r^2 < 0.30 within a 100-kb window) and maximize the instrument strength. In addition, cis-eQTL data from drug-target genes were incorporated to determine causal relationships (*Zhang et al., 2022*). Cis-eQTL were defined as eQTL within 1 Mb of the target gene. The screening criterion for eQTL effective genetic proxies was SNP <5e-08 with a minor allele frequency of >0.01.

Statistical analysis

For each exposure–outcome pair, two-sample Mendelian randomization analyses were conducted using three methods: the inverse-variance weighted (IVW) method; the weighted median method; and the Mendelian randomization–Egger method. The IVW method is the most meaningful approach; this method uses random-effects meta-analysis to amalgamate the Wald ratio estimates of the causal effect derived from each SNP. The weighted median method uses no more than half of the invalid instrument variables. The Mendelian randomization–Egger method allows all genetic variants to have pleiotropic effects, but assumes that the basic Mendelian randomization pleiotropic assumptions are not violated (*Chen et al., 2021*). Therefore, the Mendelian randomization–Egger method and the Mendelian randomization weighted median method were regarded as supplementary analyses.

In the bidirectional two-sample Mendelian randomization framework, 'lipid traits' and 'endometriosis' were interchanged as exposure and outcome to determine whether there was reverse causality between the two conditions. The false discovery rate (FDR)-adjusted *P*-value in the IVW model was estimated using the sequential *P*-value approach (*Nalls et al., 2019*). Significance was determined using an FDR-adjusted *P*-value (<0.05) in the IVW model. Instrumental variants in drug-target Mendelian randomization with weak linkage disequilibrium (r^2 < 0.3) underwent stricter linkage disequilibrium threshold tests (r^2 < 0.1, 0.01 and 0.001) to confirm robust associations.

Bayesian colocalization was employed to verify the exclusion restriction assumption, analysing eQTL data and yielding posterior probabilities of causal variants and colocalization. The main outputs were the posterior probability of different causal variants (H3), the probability of shared causal variants (H4), and the probability of colocalization conditional on the existence of the causal variant for the outcome [H4/(H3 + H4)]. The analysis yielded the posterior probability of colocalization on the presence of causal variant for the outcome (*Zuber et al., 2022*), which was the final output of interest. Further analysis was performed to determine potential mediators. Possible risk factors [i.e. interleukin factors, vascular endothelial growth factor (VEGF), body mass index and tumour necrosis factor-α)] were included in the two-step Mendelian randomization model to determine whether they altered the results.

Sensitivity analysis

The F-statistic was calculated to evaluate the strength of statistical power and the availability of instrumental variables; the statistical power was confirmed by applying mRnd to the calculations (Supplementary Tables S9 and S10). To verify the validity of the drug-target instrumental variables, coronary artery disease was chosen as a positive control in the analysis. Radial Mendelian randomization and Mendelian randomization–pleiotropy residual sum and outlier were utilized in the sensitivity analysis to identify and eliminate outliers and pleiotropy (*Lu et al., 2022*). Additionally, the Mendelian randomization–Egger intercept test was used to strengthen the detection of pleiotropy. Cochran's Q test was used to detect heterogeneity. Furthermore, leave-one-out plots were used to investigate if a single SNP was responsible for causality. The Steiger test was used to verify that the direction of causality was correct. The HEIDI test was used with SMR to distinguish pleiotropy from linkage, in which a result of $P < 0.05$ suggested possible over-linkage (*Xu et al., 2023*). All statistical analyses were performed using SMR Version 1.3.1 and R Version 4.2.2.

RESULTS

Effect of lipid traits on endometriosis

This bidirectional Mendelian randomization study elucidated the causal associations between various lipid traits and endometriosis. The effects of lipid traits on endometriosis are shown in FIGURE 1 and Supplementary Table S11. The genetically predicted elevated triglyceride concentration was causally associated with a higher incidence of all endometriosis phenotypes selected, with OR for endometriosis of 1.15 for 1 SD increase in the triglyceride concentration. Subtype analysis revealed significant associations with increased risk of endometriosis stages 1–2 (OR 1.15, 95% CI 1.04–1.28), endometriosis stages 3–4 (OR 1.14, 95% CI 1.05–1.25), ovarian endometriosis (OR 1.21, 95% CI 1.10–1.34) and pelvic peritoneum endometriosis (OR 1.17, 95% CI 1.06–1.30), all robust against FDR adjustment. Conversely, the genetically predicted elevated HDL-C concentration was causally associated with a lower risk of almost all endometriosis phenotypes, with OR for endometriosis of 0.87 for 1 SD increase in the HDL-C concentration. In the subtypes, the genetically determined elevated HDL-C concentration was significantly associated with endometriosis stages 3–4 (OR 0.88, 95% CI 0.80–0.97), ovarian endometriosis (OR 0.84, 95% CI 0.75–0.93) and pelvic peritoneum endometriosis (OR 0.84, 95% CI 0.76–0.94). The genetically proxied LDL-C concentration was causally associated with a higher incidence of endometriosis stages 1–2 and pelvic peritoneum endometriosis (OR_1 1.20, 95% CI 1.06–1.34; OR_2 1.21, 95% CI 1.06–1.37). The genetically proxied APOA concentration was causally associated with decreased risk of endometriosis (OR 0.89, 95% CI 0.83–0.96). The Steiger test confirmed correct orientation, and sensitivity analysis detected no significant pleiotropy or heterogeneity.

Effect of endometriosis on lipid traits

The effects of different types of endometriosis on lipid traits are shown in FIGURE 2 and Supplementary Table S12. Reverse directional Mendelian randomization results revealed that genetically proxied endometriosis was causally associated with an increased triglyceride concentration (OR 1.02, 95% CI 1.01–1.03). Sensitivity analysis revealed no abnormalities.

Risk of lipid-lowering drug targets and endometriosis

The effect of lipid-lowering drug targets on endometriosis is shown in FIGURE 3 and Supplementary Table S13. With FDR

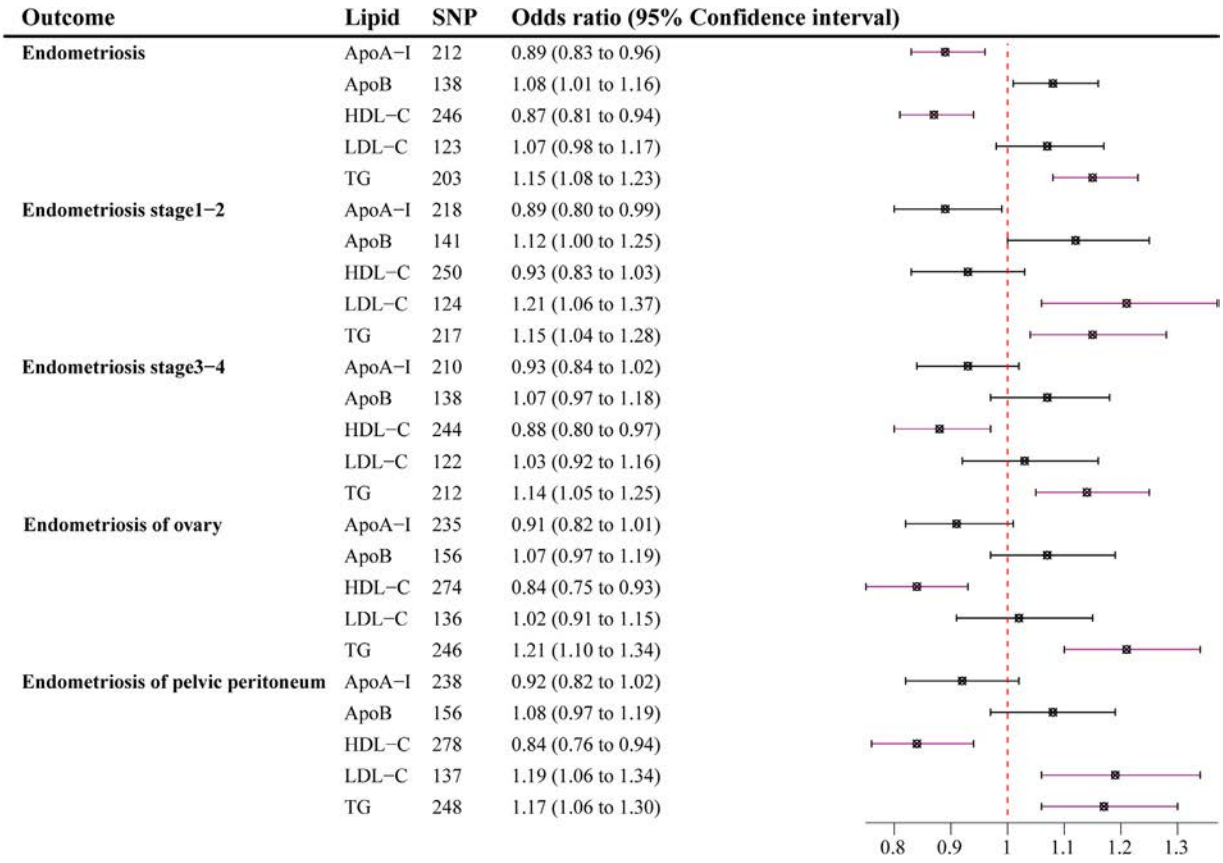


FIGURE 1 Forest plot of two-sample Mendelian randomization results. The exposure was lipid traits and outcome was endometriosis. The inverse-variance weighed analysis method was used. Exposures derived from UK Biobank and outcomes derived from the FinnGen Consortium. *P*-values that reach the false discovery rate correction threshold are marked in purple. SNP, single-nucleotide polymorphisms; ApoA-I, apolipoprotein A-I; ApoB, apolipoprotein B; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TG, triglycerides.

correction, the Mendelian randomization analysis did not provide enough evidence for the association between the drug-target gene and total endometriosis. For the subtypes, genetic mimicry in *PCSK9* and *APOB* was associated with increased risk of endometriosis stages 1–2 (*PCSK9*: OR 1.40, 95% CI 1.14–1.72; *APOB*: OR 1.50 95% CI 1.21–1.86). Genetic mimicry in *ANGPTL3* was also significantly associated with endometriosis stages 1–2 (OR 1.57 per 1-mmol/l increase in triglyceride concentration, 95% CI 1.14–2.16). For lesion locations, genetic mimicry in *PCSK9* and *APOB* was associated with the risk of pelvic peritoneum endometriosis (OR_1 1.40, 95% CI 1.13–1.73; OR_2 1.45, 95% CI 1.14–1.86). The genetic mimicry of all drug targets had neutral effects on endometriosis stages 3–4 and ovarian endometriosis. Neither pleiotropy nor heterogeneity were observed, improving the causal inferences. Most positive results remained stable even at the most stringent linkage disequilibrium threshold ($r^2 < 0.001$), except the *APOB*–pelvic peritoneum endometriosis pair (Supplementary Tables S13–S15).

The eQTL data from Genotype-Tissue Expression (GETx) V8 varied across tissues. Due to limited eQTL data for critical tissues, alternative sources were used. No *ANGPTL3* variants survived genetic instrument construction, so they were excluded from the SMR analysis. SMR supported the link between *APOB* (expressed in adipose subcutaneous tissue) and endometriosis stages 1–2 (Beta = 0.14, *P* = 0.01, with a HEIDI *P* of 0.87) (Supplementary Table S16).

Colocalization methods were performed to determine the probability of shared causal SNP for eQTL associated with gene expression in different tissues and endometriosis. The posterior probability of colocalization between LDL-C and endometriosis stages 1–2 in the *APOB* gene region was 95.7% in the presence of a causal variant for the outcome. Under the same causal conditions, the posterior probabilities for *PCSK9* and *ANGPTL3* were 88.9% and 98.7%, respectively. The posterior probability of colocalization between LDL-C and pelvic peritoneum endometriosis in the *PCSK9* gene region was 81.3% (Supplementary Table S17). Mediation analysis suggested interleukin-1α and VEGF as potential mediators in the association between APOB and endometriosis stages 1–2 (P_1 = 0.015 and P_2 = 0.026) (Supplementary Tables S18 and S19).

DISCUSSION

To the authors' knowledge, this is the first study to establish a potential bidirectional causal relationship between dyslipidaemia and endometriosis. Bidirectional Mendelian randomization analysis suggested that an elevated triglyceride concentration is a bidirectional predisposing factor for endometriosis, while an elevated LDL-C concentration appears to be a unidirectional risk factor

Outcome	Exposure	SNP	Odds ratio (95% Confidence interval)
TG	Endometriosis	20	1.02 (1.01 to 1.03)
	Endometriosis stage1−2	9	1.01 (1.00 to 1.02)
	Endometriosis stage3−4	18	1.00 (0.99 to 1.01)
	Endometriosis of ovary	20	1.00 (1.00 to 1.01)
	Endometriosis of pelvic peritoneum	10	1.01 (1.00 to 1.02)
LDL−C	Endometriosis	20	1.00 (0.99 to 1.01)
	Endometriosis stage1−2	10	1.00 (0.99 to 1.02)
	Endometriosis stage3−4	19	1.00 (0.99 to 1.01)
	Endometriosis of ovary	21	1.00 (0.99 to 1.01)
	Endometriosis of pelvic peritoneum	11	1.01 (0.99 to 1.02)
HDL−C	Endometriosis	17	0.99 (0.98 to 1.00)
	Endometriosis stage1−2	9	0.99 (0.98 to 1.00)
	Endometriosis stage3−4	18	1.00 (0.99 to 1.01)
	Endometriosis of ovary	19	0.99 (0.98 to 1.00)
	Endometriosis of pelvic peritoneum	10	0.99 (0.98 to 1.00)
ApoB	Endometriosis	20	1.00 (0.99 to 1.01)
	Endometriosis stage1−2	10	1.01 (0.99 to 1.02)
	Endometriosis stage3−4	19	1.00 (0.99 to 1.01)
	Endometriosis of ovary	21	1.00 (0.99 to 1.01)
	Endometriosis of pelvic peritoneum	11	1.01 (1.00 to 1.02)
ApoA−I	Endometriosis	19	1.00 (0.99 to 1.01)
	Endometriosis stage1−2	9	0.99 (0.98 to 1.00)
	Endometriosis stage3−4	19	1.00 (0.99 to 1.01)
	Endometriosis of ovary	20	0.99 (0.98 to 1.00)
	Endometriosis of pelvic peritoneum	10	1.00 (0.99 to 1.00)

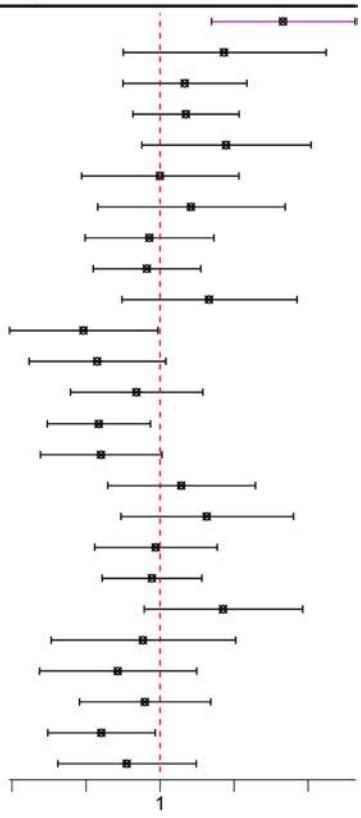


FIGURE 2 Forest plot of reverse two-sample Mendelian randomization results. Exposure was endometriosis and outcome was lipid traits. The inverse-variance weighed analysis method was used. *P*-values that reach the false discovery rate correction threshold are marked in purple. SNP, single-nucleotide polymorphisms; ApoA-I, apolipoprotein A-I; ApoB, apolipoprotein B; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TG, triglycerides.

for endometriosis. The drug-target analysis identified *PCSK9*, *APOB* and *ANGPTL3* as potential therapeutic targets for early-stage endometriosis, with *PCSK9* also relevant for pelvic peritoneum endometriosis.

This study showed intricate interplay between dyslipidaemia and endometriosis, expanding on previous research (*Gibran et al., 2017*; *Melo et al., 2010*). It established a comprehensive framework linking lipid metabolism to endometriosis, and confirmed that triglyceride and LDL-C concentrations are reliable risk factors for endometriosis. Moreover, these controllable factors provide new directions for the treatment of endometriosis. Notably, given that endometriosis appears to be more prevalent in younger, leaner women in clinical practice (*Kim et al., 2021*), there is a tendency to overlook screening for metabolic function, particularly serum lipid concentrations. Consequently, early lipid abnormalities in these women may go undetected. In light of these observations, the authors advocate for enhanced lipid metabolism screening in patients undergoing initial endometriosis evaluation. This proactive strategy is crucial to uncover hidden risk factors. For patients diagnosed concurrently with endometriosis and dyslipidaemia, the implementation of lipid-lowering therapies, supervised by an endocrinologist, may emerge as a promising dual management strategy.

Additionally, this research corroborated the identification of endometriosis as a potential risk factor for an elevated triglyceride concentration, aligning with numerous studies indicating increased risk of the subsequent development of cardiovascular disease in women with endometriosis (*Santoro et al., 2015*; *Tan et al., 2019*). Alterations in lipid metabolism, chronic inflammatory conditions, and the accumulation of oxidative stress byproducts associated with endometriosis may culminate in a gradual decline in cardiovascular health (*Lin and Li, 2022*). Therefore, the management of endometriosis should extend beyond symptom relief during a woman's reproductive years, acknowledging its profound influence on lipid metabolism and cardiovascular health in postmenopausal women.

The use of human genetic data in drug development has gained momentum in recent years (*Holmes et al., 2021*). The findings of the present study regarding *ANGPTL3* and other lipid-lowering drug targets suggest potential protective effects against endometriosis. However, the exact pharmacological mechanisms remain to be elucidated fully. The bidirectional Mendelian randomization analysis underscores the pivotal role of triglyceride concentration in the development of endometriosis. *ANGPTL3*, a central regulator of serum lipid and lipoprotein metabolism, exerts its influence by modulating lipoprotein lipase, thereby impacting the clearance of triglycerides and LDL-C (*Burks et al., 2023*). Current clinical evidence advocates for the use of ANGPTL3 inhibitors in treating hypercholesterolaemia (*Lim, 2021*). The present study supports *ANGPTL3*, which

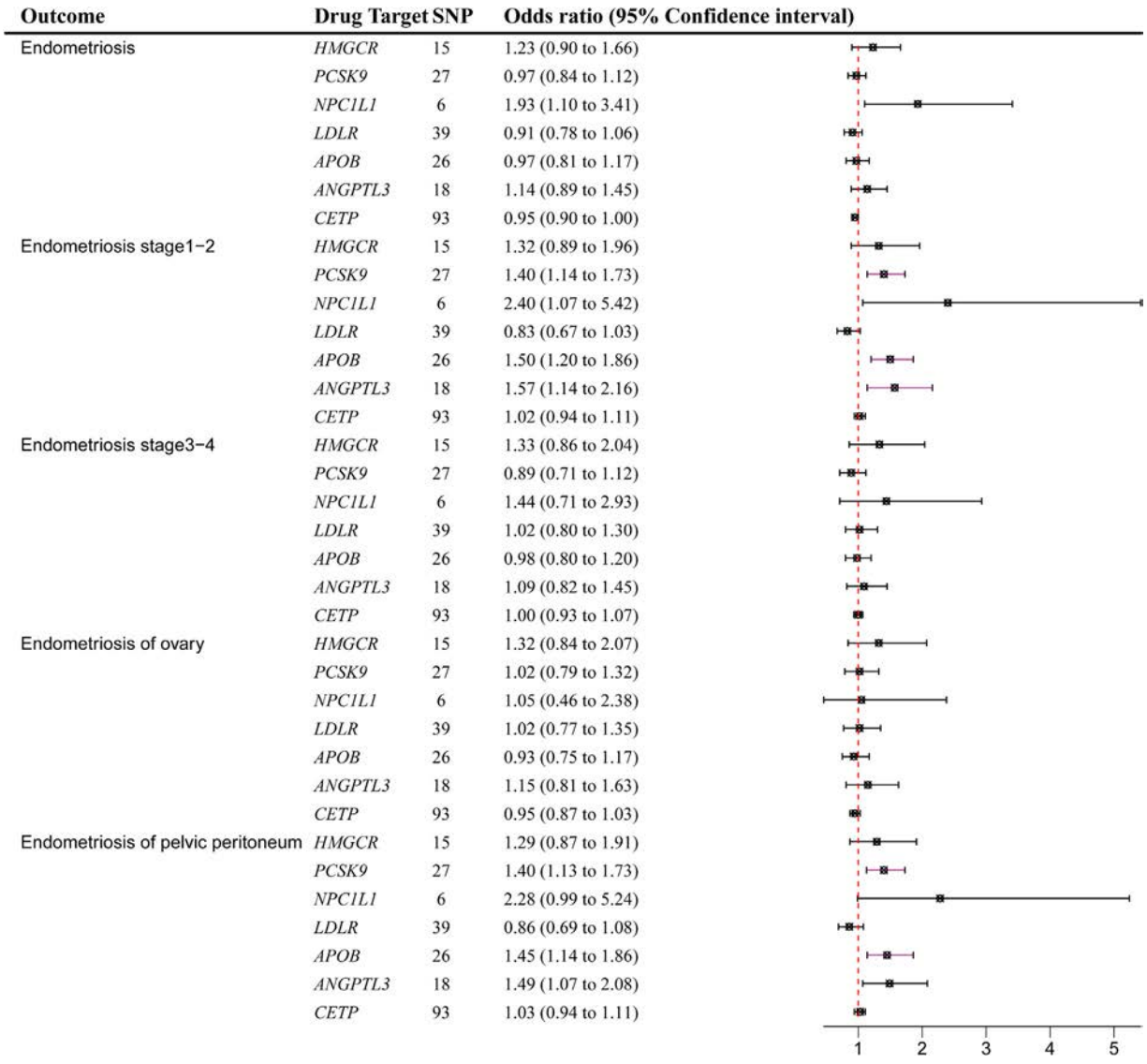


FIGURE 3 Forest plot of drug-target Mendelian randomization results. The inverse-variance weighed analysis method was used. *P*-values that reach the false discovery rate correction threshold are marked in purple. *PCSK9*, proprotein convertase subtilisin/kexin type 9; *APOB,* apolipoprotein B-100; *ANGPTL3*, angiopoietin-like 3; *HMGCR*, HMG-CoA reductase; *NPC1L1*, Niemann-Pick C1-like protein 1; *LDLR*, low-density lipoprotein receptor; *CETP*, cholesteryl ester transfer protein; SNP, single-nucleotide polymorphisms.

mediates triglyceride-lowering genes, as a potential therapeutic target in the treatment of endometriosis, presenting significant therapeutic prospects for the condition.

In the present study, only LDL-lowering genetic variants in *PCSK9* and *APOB* were associated with a lower risk of endometriosis; associations between other LDL-lowering genetic variants and endometriosis were not validated. The lack of association suggests that *PCSK9* and *APOB* genes may have physiological effects in addition to LDL metabolism (*Katsuki et al., 2022*). There is accumulating evidence suggesting that the *APOB* gene is implicated in inflammatory pathways (*Mokhtary et al., 2022*; *Song et al., 2017*). The two-step Mendelian randomization analysis undertaken in the present study showed that the reduced risk of endometriosis associated with *APOB* inhibition may be influenced, in part, by the modulation of interleukin-1 and VEGF. These inflammatory mediators play a significant role in immune evasion and the inflammatory response within endometriotic lesions. VEGF and its receptor are not only crucial for angiogenesis in eutopic endometrium, but also play an integral role in the lipid peroxidation pathway (*Ma et al., 2022*). This finding suggests that APOB inhibitors may influence the progression of endometriosis by modulating inflammatory pathways and VEGF expression. The *PCSK9* gene, known for its role in LDL-C clearance, has emerged as a particularly promising target. Recent studies have reported that PCSK9 inhibitors, traditionally used for hypercholesterolaemia, may induce ferroptosis in cells by disrupting cellular metabolism and lipid peroxidation (*Alannan et al., 2022*; *Katsuki et al., 2022*). Given the known iron overload and ferroptosis in endometriosis (*Li et al., 2021*; *Ni et al., 2022*), this mechanistic pathway warrants further exploration.

However, these studies have not determined definitively whether the dysregulation of LDL-C or triglyceride concentrations is a direct consequence of the natural pathological progression of endometriosis, or influenced by extraneous factors. The present findings underscore the imperative for more focused, mechanistic research into the role of lipid-lowering drugs in treating endometriosis and chronic inflammatory conditions. Moreover, the potential for repurposing lipid-lowering drugs in the treatment of endometriosis, as suggested by the genetic evidence in this study, necessitates further clinical corroboration through additional randomized controlled trials.

Some studies have suggested that statins could be effective for the treatment of endometriosis due to their antiangiogenic, anti-inflammatory and antiproliferative properties (*Dillon et al., 2022*; *Goenka et al., 2017*; *Sokalska et al., 2019*). However, the present study did not find a significant association between endometriosis and genetically predicted *HMGCR*. This unexpected finding may be attributed to several factors. Primarily, the methodology of the present study, rooted in genetic-level drug screening across a broad population, may lack the precision required for accurate determination of efficacy. Furthermore, previous research has been largely confined to in-vitro or animal model studies, casting uncertainty on its relevance to humans. Future investigations should integrate advanced statistical techniques and comprehensive clinicological data for thorough evaluation of the potential of statins as a treatment option for individuals with endometriosis.

While the present findings have been corroborated through various genetic variant construction frameworks, there are several limitations that merit consideration. Firstly, the predominantly European composition of the study population restricts the extrapolation of the results to non-European demographics. Secondly, a notable distinction exists between the short-term effects of pharmacological interventions and the lifelong influence of genetic variations, rendering the impact of this temporal disparity non-negligible. Thirdly, the ongoing refinement of eQTL information in GETx V8 and the absence of certain eQTL constrain the scope of the SMR analysis results. Fourthly, constraints in the initial data sets, established prior to this study, resulted in a lack of adequate statistical power for their integration into the analysis. Finally, the methodology did not account for potential off-target effects, which remains a limitation of the analyses.

CONCLUSION

The bidirectional Mendelian randomization results revealed that genetically predicted triglyceride, HDL-C, LDL-C and APOA concentrations were causally associated with endometriosis, while genetically predicted endometriosis was also causally associated with an elevated triglyceride concentration. The results contribute significantly to understanding of the interplay between lipid concentrations and endometriosis, advocating for a more integrated approach in treating this condition. In conjunction with genetic and pharmacological findings, the study results suggest that the *PCSK9*, *APOB* and *ANGPTL3* genes should be regarded as potential treatment target genes for endometriosis at an early stage. This potential application opens new avenues for clinical intervention, warranting further exploration and validation in clinical settings.

DATA AVAILABILITY

Data will be made available on request.

FUNDING

This study was supported by the National Natural Science Foundation of China (No. 82074479); the Natural Science Foundation of Jiangsu Province (No. BK20201503); Project for Scientific and Technological Development of Traditional Chinese Medicine in Jiangsu Province (2020ZX04); Open Project of Jiangsu Health Development Research Centre (JSHD2022059); Project entrusted by Jiangsu Maternal and Child Health Research Association (JSFY202205); and special funding for the training of outstanding young doctors (No. 2023QB0141). The funding sources did not have any influence on the study design, analyses or interpretation of data.

AUTHOR CONTRIBUTIONS

Conceptualization: RJ and YWC. Methodology: RJ and HQY. Data curation: NXW. Writing – original Draft: RJ and ZG. Writing – review and editing: YQ. Visualization: TXL and HRK.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.rbmo.2024.103856.

REFERENCES

Alannan, Malak, Fatrouni, Hala, Trézéguet, Véronique, Dittrich-Domergue, Franziska, Moreau, Patrick, Siegfried, Géraldine, Liet, Benjamin, Khatib, Abdel-Majid, Grosset, Christophe F., Badran, Bassam, Fayyad-Kazan, Hussein, Merched, Aksam J., 2022. Targeting PCSK9 in Liver Cancer Cells Triggers Metabolic Exhaustion and Cell Death by Ferroptosis. Cells 12 (1). https://doi.org/10.3390/cells12010062.

Bulun, Serdar E., Yilmaz, Bahar D., Sison, Christia, Miyazaki, Kaoru, Bernardi, Lia, Liu, Shimeng, Kohlmeier, Amanda, Yin, Ping, Milad, Magdy, Wei, JianJun, 2019. Endometriosis. Endocrine Reviews 40 (4), 1048–1079. https://doi.org/10.1210/er.2018-00242.

Burks, Kendall H., Basu, Debapriya, Goldberg, Ira J., Stitziel, Nathan O., 2023. Angiopoietin-like 3: An important protein in regulating lipoprotein levels. Best Practice & Research. Clinical Endocrinology & Metabolism 37 (3), 101688. https://doi.org/10.1016/j.beem.2022.101688.

Chen, Xiong, Kong, Jianqiu, Pan, Jiexin, Huang, Kai, Zhou, Wenhao, Diao, Xiayao, Cai, Jiahao, Zheng, Junjiong, Yang, Xuefan, Xie, Weibin, Yu, Hao, Li, Jiande, Pei, Lu, Dong, Wen, Qin, Haide, Huang, Jian, Lin, Tianxin, 2021. Kidney damage causally affects the brain cortical structure: A Mendelian randomization study. EBioMedicine 72, 103592. https://doi.org/10.1016/j.ebiom.2021.103592. https://pubmed.ncbi.nlm.nih.gov/34619639.

Dillon, Gabrielle A., Stanhewicz, Anna E., Serviente, Corinna, Flores, Valerie A., Stachenfeld, Nina, Alexander, Lacy M., 2022. Seven days of statin treatment improves nitric-oxide mediated endothelial-dependent cutaneous microvascular function in women with endometriosis. Microvascular Research 144, 104421. https://doi.org/10.1016/j.mvr.2022.104421.

Emdin, Connor A., Khera, Amit V., Kathiresan, Sekar, 2017. Mendelian Randomization. JAMA 318 (19), 1925–1926. https://doi.org/10.1001/jama.2017.17219.

Gibran, Luciano, Maranhão, Raul C., Tavares, Elaine R., Carvalho, Priscila O., Abrão, Maurício S., Podgaec, Sergio, 2017. mRNA levels of low-density lipoprotein receptors are overexpressed in the foci of deep bowel endometriosis. Human Reproduction (Oxford, England) 32 (2), 332–339. https://doi.org/10.1093/humrep/dew303.

Goenka, Luxitaa, George, Melvin, Sen, Maitrayee, 2017. A peek into the drug development scenario of endometriosis - A systematic review. Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie 90, 575–585. https://doi.org/10.1016/j.biopha.2017.03.092.

Guo, Sun-Wei., 2023. Various types of adenomyosis and endometriosis: in search of optimal management. Fertility and Sterility 119 (5), 711–726 . https://doi.org/10.1016/j.fertnstert.2023.03.021.

Holmes, Michael V., Richardson, Tom G., Ference, Brian A., Davies, Neil M., Smith, George Davey, 2021. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development.'. Nature Reviews. Cardiology 18 (6), 435–453. https://doi.org/10.1038/s41569-020-00493-1.

Huang, Wuqing, Xiao, Jun, Ji, Jianguang, Chen, Liangwan, 2021. Association of lipid-lowering drugs with COVID-19 outcomes from a Mendelian randomization study. ELife 10. https://doi.org/10.7554/eLife.73873.

Katsuki, Shunsuke, Jha, Prabhash K, Lupieri, Adrien, Nakano, Toshiaki, Passos, Livia S.A., Rogers, Maximillian A., Becker-Greene, Dakota, Le, Thanh-Dat, Decano, Julius L., Lee, Lang Ho, Guimaraes, Gabriel C., Abdelhamid, Ilyes, Halu, Arda, Muscoloni, Alessandro, Cannistraci, Carlo V, Higashi, Hideyuki, Zhang, Hengmin, Vromman, Amélie, Libby, Peter, Keith Ozaki, C., Sharma, Amitabh, Singh, Sasha A., Aikawa, Elena, Aikawa, Masanori, 2022. Proprotein Convertase Subtilisin/Kexin 9 (PCSK9) Promotes Macrophage Activation via LDL Receptor-Independent Mechanisms. Circulation Research 131 (11), 873–889. https://doi.org/10.1161/CIRCRESAHA.121.320056.

Kim, H.J., Lee, H.S., Kazmi, S.Z., Hann, H.J., Kang, T., Cha, J., Choi, S., Swan, H., Kim, H., Lee, Y.S., Ahn, H.S., 2021. Familial risk for endometriosis and its interaction with smoking, age at menarche and body mass index: a population-based cohort study among siblings. BJOG: an International Journal of Obstetrics and Gynaecology 128 (12), 1938–1948. https://doi.org/10.1111/1471-0528.16769.

Li, Yajie, Zeng, Xinliu, Lu, Dingheng, Yin, Minuo, Shan, Meirong, Gao, Ying, 2021. Erastin induces ferroptosis via ferroportin-mediated iron accumulation in endometriosis. Human Reproduction (Oxford, England) 36 (4), 951–964. https://doi.org/10.1093/humrep/deaa363.

Liang, Zongwen, Wu, Qiong, Wang, Honglin, Tan, Jiahuan, Wang, Han, Gou, Yanling, Cao, Yingying, Li, Zhi, Zhang, Zongfeng, 2022. Silencing of lncRNA MALAT1 facilitates erastin-induced ferroptosis in endometriosis through miR-145-5p/MUC1 signaling. Cell Death Discovery 8 (1), 190. https://doi.org/10.1038/s41420-022-00975-w.

Lim, Gregory B., 2021. ANGPTL3 inhibition for hypercholesterolaemia.'. Nature Reviews. Cardiology 18 (2), 72. https://doi.org/10.1038/s41569-020-00483-3.

Lin, Xiao-Mei, Li, Dong-Zhi, 2022. Assessment of atherosclerosis in endometriosis patients: the need to do much more. American Journal of Obstetrics and Gynecology 227 (4), 670–671. https://doi.org/10.1016/j.ajog.2022.06.036.

Lu, Yao, Tang, Haibo, Huang, Peiyuan, Wang, Jie, Deng, Peizhi, Li, Yalan, Zheng, Jie, Weng, Liang, 2022. Assessment of causal effects of visceral adipose tissue on risk of cancers: a Mendelian randomization study. International Journal of Epidemiology 51 (4), 1204–1218. https://doi.org/10.1093/ije/dyac025.

Ma, Caiqi, Huang, Wei, Wang, Hui, Yao, Wenxia, Liang, Min, Yu, Guifang, Zhou, Xinke, 2022. Oxidized LDL promotes EMS-induced angiogenesis by increasing VEGF-A expression and secretion by endometrial cells. Molecular Medicine (Cambridge, Mass.) 28 (1), 151. https://doi.org/10.1186/s10020-022-00582-6.

Melo, Anderson Sanches, Rosa-e-Silva, Julio César, Rosa-e-Silva, Ana Carolina Japur de Sá, Poli-Neto, Omero Benedicto, Ferriani, Rui Alberto, Vieira, Carolina Sales, 2010. Unfavorable lipid profile in women with endometriosis. Fertility and Sterility 93 (7), 2433–2436. https://doi.org/10.1016/j.fertnstert.2009.08.043.

Mokhtary, Nasim, Neda Mousavi, Seyedeh, Sotoudeh, Gity, Qorbani, Mostafa, Dehghani, Maryam, Koohdani, Fariba, 2022. Deletion allele of Apo B gene is associated with higher inflammation, oxidative stress and dyslipidemia in obese type 2 diabetic patients: an analytical cross-sectional study. BMC Endocrine Disorders 22 (1), 73. https://doi.org/10.1186/s12902-022-00991-y.

Mu, Fan, Rich-Edwards, Janet, Rimm, Eric B., Spiegelman, Donna, Forman, John P., Missmer, Stacey A., 2017. Association Between Endometriosis and Hypercholesterolemia or Hypertension. Hypertension (Dallas, Tex.: 1979) 70 (1), 59–65. https://doi.org/10.1161/HYPERTENSIONAHA.117.09056.

Nalls, Mike A., Blauwendraat, Cornelis, Vallerga, Costanza L., Heilbron, Karl, Bandres-Ciga, Sara, Chang, Diana, Tan, Manuela, Kia, Demis A., Noyce, Alastair J., Xue, Angli, Bras, Jose, Young, Emily, Coelln, Rainer von, Simón-Sánchez, Javier, Schulte, Claudia, Sharma, Manu, Krohn, Lynne, Pihlstrøm, Lasse, Siitonen, Ari, Iwaki, Hirotaka, Leonard, Hampton, Faghri, Faraz, Raphael Gibbs, J., Hernandez, Dena G., Scholz, Sonja W., Botia, Juan A., Martinez, Maria, Corvol, Jean-Christophe, Lesage, Suzanne, Jankovic, Joseph, Shulman, Lisa M., Sutherland, Margaret, Tienari, Pentti, Majamaa, Kari, Toft, Mathias, Andreassen, Ole A., Bangale, Tushar, Brice, Alexis, Yang, Jian, Gan-Or, Ziv, Gasser, Thomas, Heutink, Peter, Shulman, Joshua M., Wood, Nicholas W., Hinds, David A., Hardy, John A., Morris, Huw R., Gratten, Jacob, Visscher, Peter M., Graham, Robert R., Singleton, Andrew B., 2019. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. The Lancet. Neurology 18 (12), 1091–1102. https://doi.org/10.1016/S1474-4422(19)30320-5.

Ni, Zhexin, Li, Yangshuo, Song, Di, Ding, Jie, Mei, Shanshan, Sun, Shuai, Cheng, Wen, Yu, Jin, Zhou, Ling, Kuang, Yanping, Li, Mingqing, Cai, Zailong, Yu, Chaoqin, 2022. Iron-overloaded follicular fluid increases the risk of endometriosis-related infertility by triggering granulosa cell ferroptosis and oocyte dysmaturity. Cell Death & Disease 13 (7), 579. https://doi.org/10.1038/s41419-022-05037-8.

Pahan, K., 2006. Lipid-lowering drugs. Cellular and Molecular Life Sciences: CMLS 63 (10), 1165–1178 .

Santoro, Luca, D'Onofrio, Ferruccio, Flore, Roberto, Gasbarrini, Antonio, Santoliquido, Angelo, 2015. Endometriosis and atherosclerosis: what we already know and what we have yet to discover. American Journal of Obstetrics and Gynecology 213 (3), 326–331. https://doi.org/10.1016/j.ajog.2015.04.027.

Skrivankova, Veronika W., Richmond, Rebecca C., Woolf, Benjamin A.R., Yarmolinsky, James, Davies, Neil M., Swanson, Sonja A., VanderWeele, Tyler J., Higgins, Julian P.T., Timpson, Nicholas J., Dimou, Niki, Langenberg, Claudia, Golub, Robert M., Loder, Elizabeth W., Gallo, Valentina, Tybjaerg-Hansen, Anne, Smith, George Davey, Egger, Matthias, Brent Richards, J., 2021. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. JAMA 326 (16), 1614–1621. https://doi.org/10.1001/jama.2021.18236.

Sokalska, Anna, Hawkins, Amanda B., Yamaguchi, Toshia, Duleba, Antoni J., 2019. Lipophilic statins inhibit growth and reduce invasiveness of human endometrial stromal cells.

Journal of Assisted Reproduction and Genetics 36 (3), 535–541. https://doi.org/10.1007/s10815-018-1352-9.

Song, Yongyan, Zhao, Miaoyun, Cheng, Xiao, Shen, Jing, Khound, Rituraj, Zhang, Kezhong, Su, Qiaozhu, 2017. CREBH mediates metabolic inflammation to hepatic VLDL overproduction and hyperlipoproteinemia. Journal of Molecular Medicine (Berlin, Germany) 95 (8), 839–849. https://doi.org/10.1007/s00109-017-1534-4.

Tan, Justin, Taskin, Omur, Iews, Mahmoud, Lee, Arthur J., Kan, Arohumam, Rowe, Timothy, Bedaiwy, Mohamed A., 2019. Atherosclerotic cardiovascular disease in women with endometriosis: a systematic review of risk factors and prospects for early surveillance. Reproductive Biomedicine Online 39 (6), 1007–1016. https://doi.org/10.1016/j.rbmo.2019.05.021.

Taylor, Hugh S., Iii, Myles Alderman, D'Hooghe, Thomas M., Fazleabas, Asgerally T., Duleba, Antoni J., 2017. Effect of simvastatin on baboon endometriosis. Biology of Reproduction 97 (1), 32–38. https://doi.org/10.1093/biolre/iox058.

Vercellini, Paolo, Viganò, Paola, Somigliana, Edgardo, Fedele, Luigi, 2014. Endometriosis: pathogenesis and treatment. Nature Reviews. Endocrinology 10 (5), 261–275. https://doi.org/10.1038/nrendo.2013.255.

Xu, Shu, Li, Xiaozhi, Zhang, Shenghong, Qi, Cancan, Zhang, Zhenhua, Ma, Ruiqi, Xiang, Liyuan, Chen, Lianmin, Zhu, Yijun, Tang, Ce, Bourgonje, Arno R., Li, Miaoxin, He, Yao, Zeng, Zhirong, Hu, Shixian, Feng, Rui, Chen, Minhu, 2023. Oxidative stress gene expression, DNA methylation, and gut microbiota interaction trigger Crohn's disease: a multi-omics Mendelian randomization study. BMC Medicine 21 (1), 179. https://doi.org/10.1186/s12916-023-02878-8.

Zhang, Xiaoyu, Geng, Tao, Li, Ning, Wu, Lijuan, Wang, Youxin, Zheng, Deqiang, Guo, Bo, Wang, Baoguo, 2022. Associations of Lipids and Lipid-Lowering Drugs with Risk of Vascular Dementia: A Mendelian Randomization Study. Nutrients 15 (1). https://doi.org/10.3390/nu15010069.

Zondervan, Krina T., Becker, Christian M., Missmer, Stacey A., 2020. 'Endometriosis. The New England Journal of Medicine 382 (13), 1244–1256. https://doi.org/10.1056/NEJMra1810764.

Zuber, Verena, Grinberg, Nastasiya F., Gill, Dipender, Manipur, Ichcha, Slob, Eric A.W., Patel, Ashish, Wallace, Chris, Burgess, Stephen, 2022. Combining evidence from Mendelian randomization and colocalization: Review and comparison of approaches. American Journal of Human Genetics 109 (5), 767–782. https://doi.org/10.1016/j.ajhg.2022.04.001.

Received 10 October 2023; received in revised form 25 January 2024; accepted 26 January 2024.

RBMO

ELSEVIER

ARTICLE

Cardiovascular disease hospitalizations among women who undergo fertility treatment

BIOGRAPHY

Dr Dee Tomic is a clinically trained postdoctoral research fellow at Monash University School of Public Health and Preventive Medicine. She completed her PhD in diabetes epidemiology with the Baker Heart and Diabetes Institute in 2023. Dee is skilled in working with clinical registries, cohort studies, hospital and mortality data.

Dunya Tomic[a,*], Luk Rombauts[b,c], Fabricio Da Silva Costa[d,e], Daniel L. Rolnik[a,b], Tilahun Haregu[f], Melinda Carrington[f], Dianna J. Magliano[a,f,1], Stephanie R. Yiallourou[a,g,1]

KEY MESSAGE

In this study of 27,262 Australian women, those who received fertility treatment were not at increased risk of cardiovascular disease hospitalization compared with those who did not. In conjunction with similar results from other studies, this finding provides reassurance that fertility therapy is a safe option for many women.

ABSTRACT

Research question: Are women who receive fertility treatment at increased risk of cardiovascular disease (CVD) hospitalization compared with women who do not?

Design: A retrospective cohort study of all women registered for fertility treatment at Monash IVF between 1998 and 2014. This cohort was linked to the Victorian Admitted Episodes Dataset, which contains records of all hospital admissions in the Australian state of Victoria. Age- and Index of Relative Socioeconomic Disadvantage (IRSD)-adjusted relative risks of CVD hospitalization for women who did or did not undergo fertility treatment were determined using Poisson regression. Risks were calculated overall by CVD subtype and stratified by area-based social disadvantage using IRSD fifths, number of stimulated cycles and mean oocytes per cycle.

Results: Of 27,262 women registered for fertility treatment, 24,131 underwent treatment and 3131 did not. No significant difference was found in risk of CVD hospitalization between treated and untreated women overall (adjusted RR 0.93, 95% 0.82 to 1.05) or by CVD subtype. The admission risk for CVD was significantly lower in treated women who had a mean of fewer than five oocytes per cycle (adjusted RR 0.80, 95% CI 0.70 to 0.92) compared with untreated women. Treated women residing in areas within the second IRSD fifth were less likely to be hospitalized for CVD compared with untreated women (age-adjusted RR 0.66, 95% CI 0.49 to 0.89).

Conclusions: Fertility treatment is not associated with increased risk of CVD hospitalization. Lower risk among some subgroups of treated women may be explained by social disadvantage.

[a] Monash University, Melbourne, Australia.
[b] Monash Health, Melbourne, Australia.
[c] Monash IVF, Melbourne, Australia.
[d] Gold Coast University Hospital, Gold Coast, Australia.
[e] Griffith University, Gold Coast, Australia.
[f] Baker Heart and Diabetes Institute, Melbourne, Australia.
[g] The Turner Institute for Brain and Mental Health, Monash University, Melbourne, Australia.
[1] Joint first authors.

KEYWORDS
Cardiovascular disease
Assisted reproductive technologies
Fertility
Hospitalisation
Registries

*Corresponding author. E-mail address: Dee.Tomic@monash.edu (D. Tomic).
https://doi.org/10.1016/j.rbmo.2024.103812
1472-6483/
Declaration: Monash IVF supported this investigation through the Monash Research and Education Fund. Luk Rombauts was employed at Monash IVF during the period covered by the study. Daniel L Rolnik was employed by Monash Ultrasound for Women, a branch of Monash IVF.

INTRODUCTION

Assisted reproduction technology has advanced greatly since the first successful IVF birth in 1978 (*Eskew* et al., *2017*). Today, over half a million IVF deliveries occur annually worldwide (*Patrizio* et al., *2022*), and, in Australia, the proportion of babies born via IVF is now almost one in 20 (*University of New South Wales, 2020*). Ovarian stimulation is generally considered a safe option for many women; however, whether such fertility treatment is associated with adverse cardiovascular outcomes in the long term remains unclear. Fertility treatment may lead to adverse cardiovascular events by inducing background thrombosis, activating the renin-angiotensin system or inducing vascular injury from ovarian hyperstimulation (*Magnus* et al., *2023*). Indirectly, fertility treatment is associated with increased risks of pregnancy-related complications, such as maternal metabolic syndromes, e.g. gestational diabetes and hypertension, which are associated with higher risk of cardiovascular disease (CVD) in the mother and offspring in the long term (*Fraser* et al., *2012*). It is also possible that conditions that lead to infertility and subsequent fertility treatment, such as polycystic ovary syndrome, endometriosis or endometrial dysfunction, are also associated with increased risk of CVD (*Hart* et al., *2015*; *Tan* et al., *2019*; *Zhang* et al., *2022*). A Swedish population-based study found higher rates of hypertension and a trend toward more incident strokes among women who received fertility treatment compared with those who did not (*Westerlund* et al., *2014*). A systematic review of six studies and over 1.4 million women, however, found no increased risk of cardiac events associated with fertility treatment (*Dayan* et al., *2017*). More recently, our cohort study of Australian women found no increase in CVD mortality risk among women who underwent fertility treatment compared with women who were registered for, but did not undergo, fertility treatment (*Yiallourou* et al., *2022*).

Most studies that have investigated CVD risk associated with fertility treatment have focused on mortality, or a composite of mortality and other CVD events, as the major outcome measure (*Venn* et al., *2001*; *Udell* et al., *2013*; *Yiallourou* et al., *2022*). As most of the women involved in these studies are relatively young at the time of fertility treatment, most aged between 30 and 40 years (*Sneed* et al., *2008*), follow-up may have been insufficient to capture many deaths. As such, observing CVD-related hospitalizations may provide a better opportunity to assess CVD risk in this context. One cohort study of 99,291 women in Israel followed up for a mean of 11.7 years found no increased risk for CVD hospitalization among those who underwent fertility treatment (*Ben-Yaakov* et al., *2016*). These results, however, have not been replicated in other jurisdictions, and no Australian data have been published on this important topic. Furthermore, no study has looked at hospitalization risk for specific forms of CVD, e.g. hypertension, heart failure, or whether this risk varies according to factors such as social disadvantage or the amount of treatment received. As such, the aim of the present study was to determine the risk of CVD hospitalizations, overall and by subtype of CVD, and stratified by area-based social disadvantage, stimulated cycles and oocytes per cycle, among women registered with Monash IVF services in Australia who either proceeded to fertility treatment or did not.

MATERIALS AND METHODS

A retrospective cohort study that analysed clinical records in the Monash IVF clinical registry (1998–2014) was conducted. Hospitalizations for CVD among women who underwent fertility treatment were determined through data linkage to the Victorian Admitted Episodes Dataset (VAED). Data were also linked to the National Death Index (NDI) to determine any deaths that occurred among the cohort during the study period.

Study population

Data were initially sourced for all women who registered for fertility treatment at Monash IVF clinics between 1 January 1998 and 1 January 2014. The start date was selected as this was the earliest date for which VAED data were available. Women whose usual place of residence was outside Australia and those with inadequate information available for data linkage were excluded. Women were classified as 'treated' or 'untreated' depending on whether they underwent treatment with fertility drugs. Those for whom at least one ovarian stimulation to induce multiple folliculogenesis was recorded were classified as 'treated'. Those registered for fertility treatment who did not undertake treatment or had natural cycle treatment without ovarian stimulation were classified as 'untreated'. Reasons for not proceeding with treatment included pursuit of other treatment options, natural conception while waiting to be seen for treatment, and financial and relationship difficulties. As this information was not routinely collected, we could not provide a more detailed breakdown. The rationale for using women who registered for, but did not undergo, treatment as a control group was that any possible associations of fertility therapy with health outcomes may not be attributable to the treatment itself, but rather driven by disease processes that increase the risk of infertility as well as other health outcomes. From this perspective, comparing to an otherwise healthy population would be inappropriate.

Data linkage

Demographic data that were extracted for each woman included name, date of birth, residential postcode and date of registration with the clinic. Fertility treatment information that was collected included cause of infertility, number of stimulated cycles and number of oocytes collected. Cohort data were linked to VAED records from 1998 to 2019 by the Centre for Victorian Data Linkage (CVDL). The VAED is an administrative database that captures all public and private hospitalizations in the Australian state of Victoria. It contains diagnostic codes for primary and subsequent diagnoses according to the International Classification of Diseases, tenth revision (ICD-10). Up to 50 ICD-10 codes were collected for each admission. Other information pertinent to this analysis included admission and discharge dates. Linkage between the Monash IVF database and the VAED was based on full name, sex, date of birth and postcode information. The CVDL data linkage process, which involved constructing a system of links by combining personal identifier variables from 27 Victorian health and human services datasets, is presented in **Supplementary Figure 1**. Of the Monash IVF cohort, 30.4% (8298/27,262) had at least one presentation in VAED datasets. Cohort data were also linked to the National Death Index through the process described in our previous work on mortality (*Yiallourou* et al., *2022*).

Statistical analysis

All treated and untreated women were followed up from their date of last contact with Monash IVF until the end of the study period (31 December 2019), or date of death if they died during the study period. Any CVD hospitalizations that occurred during this period were determined, which were defined as those with any of the following ICD-10 codes: I10-I15, I20-I25 (ischaemic heart diseases), I46.1 (sudden cardiac death), I48 (atrial fibrillation and flutter), or RR) and 95% confidence intervals of CVD admission among treated and untreated women were estimated using Poisson regression. This was done overall and stratified by stimulated cycles in treated women (categorized as fewer than three, or three or more), mean numbers of oocytes per stimulated cycle in treated women (fewer than five, five or more), and area-based social disadvantage. Area-based social disadvantage was explored by mapping postcode data to Index of Relative Socioeconomic Disadvantage (IRSD) scores, which summarize information about social disadvantage, e.g. occupations, income levels, of people in a geographic area (*Australian Bureau of Statistics, 2022*). Lower IRSD scores indicate greater disadvantage, whereas higher scores indicate relative lack of disadvantage. The IRSD scores were split into fifths. The CVD admissions were further divided by subtype at the ICD-10 three-digit diagnosis level. Relative risks of admission and 95% confidence intervals were calculated for any diagnoses with at least 10 admissions in each group (treated and untreated) over the course of the study period. For all measures other than the analysis stratified by IRSD, crude and age- and IRSD-adjusted relative risks were calculated. For the IRSD-stratified analysis, crude and age-adjusted RRs were calculated. StataCorp. 2019. (Stata Statistical Software: Release 16, College Station, TX: StataCorp LLC.) was used for all analyses.

Ethics approval

The Monash Health Research Ethics Committee (LNR/17/MonH/385, 15 September 2017) and the Australian Institute of Health and Welfare Ethics Committee (EO2018/2/447, 17 July 2018) approved this study. The requirement for individual consent was waived.

Results

The present study included 27,262 women who were registered for fertility treatment at Monash IVF clinics between 1 January 1998 and 1 January 2014. Of these women, 24,131 subsequently underwent treatment, whereas 3131 did not. At entry, the age distribution of the two treatment status groups was similar: treated women had a median age of 35 years (interquartile range [IQR] 31–39), whereas untreated women had a median age of 36 years (interquartile range [IQR] 31–40) (TABLE 1). At follow-up, over one-half of treated (52.8%) and untreated (50.1%) women were aged between 40 and 49 years. The largest proportion (28.3%) of treated women resided in areas of least social disadvantage (fifth IRSD fifth), whereas the largest proportion (27.8%) of untreated women resided in the fourth IRSD fifth. Just over one-half (50.9%) of treated women had

TABLE 1 CHARACTERISTICS OF WOMEN REGISTERED FOR FERTILITY TREATMENT AT MONASH IVF, 1 JANUARY 1998 TO 1 JANUARY 2014

Characteristics	Treated	Untreated
Number of women	24,131	3131
Age at entry, median, years	35 (31–39)	36 (31–40)
Age at follow-up, median, years	48 (43–52)	48 (43–52)
<40	2960 (12.3)	452 (14.4)
40–49	12,747 (52.8)	1569 (50.1)
50–59	7753 (32.1)	1000 (31.9)
≥60	671 (2.8)	110 (3.5)
IRSD fifth		
1 (most disadvantaged)	2433 (10.1)	382 (12.2)
2	3115 (12.9)	418 (13.4)
3	5025 (20.8)	633 (20.2)
4	6736 (27.9)	869 (27.8)
5 (least disadvantaged)	6822 (28.3)	829 (26.5)
Stimulated cycles, median	2 (1–4)	–
1 or 2	12,292 (50.9)	–
3 or more	11,839 (49.1)	–
Oocytes per cycle, median	6.9 (3–11)	
<5	8533 (35.4)	–
5 or more	15,548 (64.4)	–
Missing data	50 (0.2)	–
Fertility drug type[a]		
FSH	23,517 (97.5)	–
Hormone replacement therapy	5632 (23.3)	–
Clomiphene citrate	1508 (6.2)	–
Unknown	8 (0.0)	–
Cause of infertility		
Cause recorded	15,358 (63.6)	1687 (53.9)
Unexplained	4788 (31.2)	1342 (79.6)
Tubal	1954 (12.7)	39 (2.3)
Endometrial	1458 (9.5)	27 (1.6)
Ovarian	1657 (10.8)	34 (2.0)
Multiple factors	2901 (18.9)	127 (7.5)
Male factor	1181 (7.7)	35 (2.1)
Other	1419 (9.2)	83 (4.9)
No aetiology recorded	8773 (36.4)	1444 (46.1)

Data presented as *n* (%) and interquartile range.

[a] It is possible that women who have previously been treated with a combination of drugs can receive multiple treatments. IRSD, Index of Relative Socioeconomic Disadvantage; IQR, interquartile range; –, Not applicable.

fewer than three stimulated cycles, whereas just under one-half (49.1%) had three or more stimulated cycles. Almost two-thirds (64.4%) of treated women had an average of five or more oocytes per stimulated cycle. The most frequently used fertility drug type was FSH, which was used by 97.5% of treated women, whereas smaller proportions of treated women used hormone replacement therapy (23.3%) and clomiphene citrate (6.2%). A cause of infertility was recorded for 63.6% of treated and 53.9% of untreated women. The most recorded cause was 'unexplained' in both treated (31.2%) and untreated women (79.6%); multiple factors were implicated in a notable proportion of treated women (18.9%).

Cardiovascular disease hospitalizations

A total of 6268 CVD admissions occurred among treated women and 523 among untreated women by the end of the follow-up period on 31 December 2019 (**Supplementary Table 1**). The median follow-up time was 11.8 years. No significant difference was found between treatment status groups in the overall risk of CVD hospitalization (adjusted RR 0.93, 95% 0.82 to 1.05) (FIGURE 1A). When stratified by area-based social disadvantage, treated women who resided in areas within the first, third, fourth and fifth IRSD fifths had no difference in CVD hospitalization risk compared with treated women (FIGURE 1B). Treated women who resided in areas within the second IRSD fifth, i.e. second most socially disadvantaged areas, had a lower risk of CVD hospitalization compared with untreated women (age-adjusted RR 0.66, 95% CI 0.49 to 0.89). When risk was stratified by number of stimulated cycles, treated women who had one or two stimulated cycles had no increased risk of CVD admission compared with untreated women (adjusted RR 0.98, 95% CI 0.86 to 1.12), whereas women who had three or more

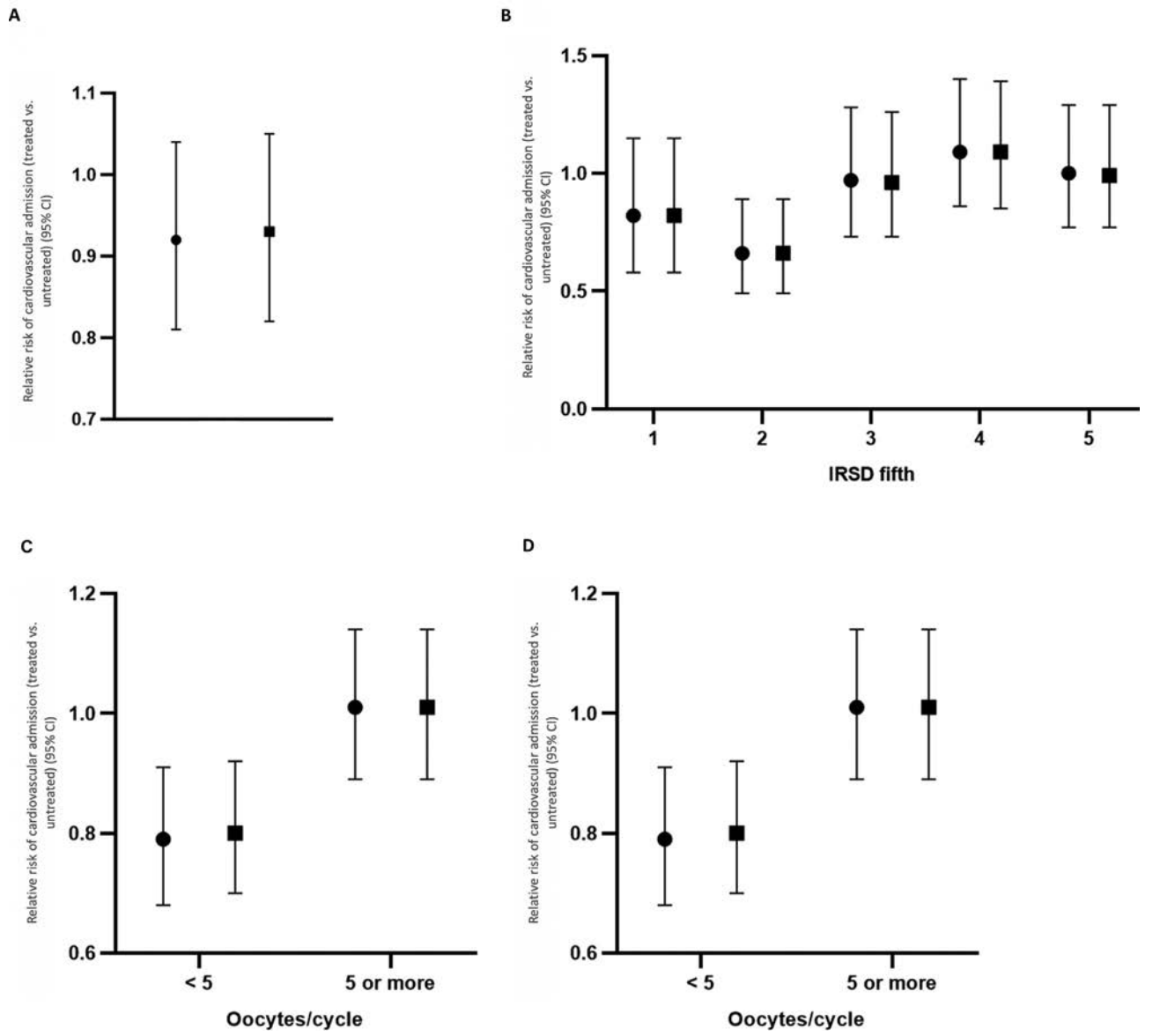


FIGURE 1 Relative risks (crude = •; age- and Index of Relative Socioeconomic Disadvantage (IRSD)-adjusted = ■) of cardiovascular hospitalization, (A) overall and by (B) IRSD, (C) stimulated cycles and (D) oocytes per cycle, for treated versus untreated women registered at Monash IVF, 1 January 1998 to 1 January 2014.

stimulated cycles had a reduced risk (adjusted RR 0.88, 95% CI 0.77 to 1.00), which failed to reach statistical significance (FIGURE 1C). When CVD admission risk was stratified by mean oocytes per cycle, no difference was found between treated women who had a mean of five or more oocytes per cycle and untreated women (adjusted RR 1.01, 95% CI 0.89 to 1.14) (FIGURE 1D). The risk of CVD admission, however, was significantly lower in treated women who had a mean of fewer than five oocytes per cycle compared with untreated women (adjusted RR 0.80, 95% CI 0.70 to 0.92). Within the treated group, women with a mean of five or more oocytes per cycle were significantly more likely to be hospitalized for CVD compared with treated women with a mean of fewer than five oocytes per cycle (adjusted RR 1.22, 95% CI 1.10 to 1.34) (FIGURE 2). Among treated women, an increase of one mean oocyte per cycle was associated with a 1% increased risk of CVD admission, and this result was borderline in terms of significance (adjusted RR 1.01, 95% CI 1.00 to 1.02).

Cardiovascular disease hospitalizations by subtype

The five most common reasons for CVD hospitalization at the ICD-10 three-digit diagnosis level in this cohort were hypertension, hypotension, chronic ischaemic heart disease, heart failure and angina pectoris (**Supplementary Table 1**). No significant difference was found in the risk of admission for any of these conditions between treated and untreated women (FIGURE 3).

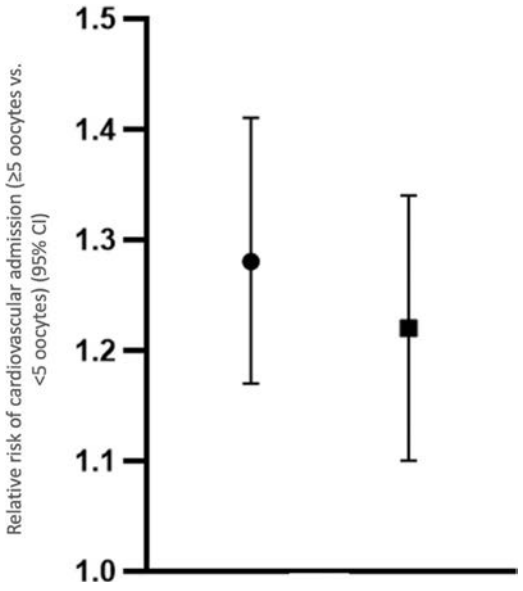


FIGURE 2 Relative risk (crude = •; age- and Index of Relative Socioeconomic Disadvantage-adjusted = ■) of cardiovascular hospitalization for treated women registered at Monash IVF with a mean of five or more oocytes per cycle compared with less than five oocytes per cycle, 1 January 1998 to 1 January 2014.

DISCUSSION

This study identified no increase in the overall risk of CVD hospitalization among women who were registered for, and underwent, fertility treatment compared with those who were registered but did not receive treatment. No differences in admissions were identified when data were stratified for the five most common CVD subtypes. Area-based social disadvantage was associated with CVD hospitalization, with reduced risk among treated versus untreated women in those with higher levels of disadvantage. Furthermore, risk of CVD admissions in treated women who averaged fewer than five oocytes per cycle was lower compared with untreated women. Overall, no harm associated with fertility treatment in terms of CVD hospitalization risk could be identified.

We extend the findings of our previous work, which found similar risk of all-cause mortality and lower risk of CVD mortality among women who received fertility treatment at Monash IVF clinics between 1982 and 2014 compared with those who were registered but did not receive treatment (*Yiallourou et al., 2022*). These results also confirm previous findings from an Israeli study (*Ben-Yaakov et al., 2016*), which reported a hazard ratio of 1.1 (95% CI 0.9 to 1.3) for CVD hospitalization among women exposed to fertility treatment compared with those who were unexposed. This Israeli study adjusted for conditions, including preeclampsia, obesity

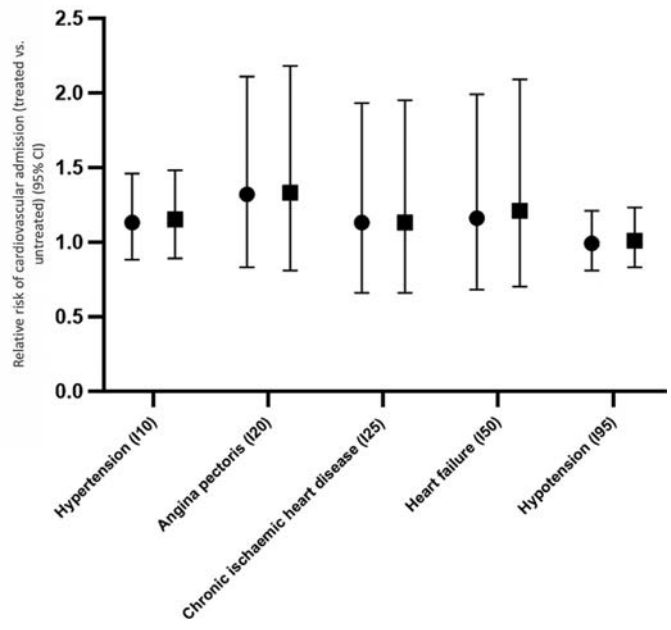


FIGURE 3 Relative risks (crude = •; age- and Index of Relative Socioeconomic Disadvantage-adjusted = ■) of cardiovascular hospitalization by the International Classification of Diseases, Tenth Revision subtype for treated versus untreated women registered at Monash IVF, 1 January 1998 to 1 January 2014.

and diabetes mellitus; however, we were unable to account for these factors. Some of these conditions, however, are unlikely to be confounding factors, but may serve as mediators, as they typically occur after fertility treatment exposure. Our study also provides further information about the specific subtypes of CVD that constitute the bulk of these hospitalizations and how the risk may be affected by factors such as area-based social disadvantage, stimulated cycles and oocytes per cycle. In contrast to our study, evidence from a population-based cohort study (*Udell* et al., *2013*) found an increased risk of long-term adverse CVD events in women whose fertility treatment was unsuccessful compared with those who successfully delivered. As we did not investigate treatment failure, our results are not directly comparable to those arising from this analysis. Women whose treatment is unsuccessful may represent a high-risk group for CVD outcomes, which should be considered in clinical advice to those prone to treatment failure.

The relative risk of CVD hospitalization between treated and untreated women was lowest in the two lowest IRSD fifths (reflecting lowest area-based social disadvantage), reaching statistical significance in the second fifth, i.e. women residing in areas of second-lowest social advantage. The relative risk increased toward the highest levels of social advantage (fourth and fifth IRSD fifths). As the 95% CI for each of the IRSD fifths overlapped, however, it cannot be concluded that there was a difference in relative risk according to IRSD. Our previous study of CVD mortality found that treated women had lower risk of CVD mortality compared with untreated women, and that the difference in risk was greatest in women of higher area-based social disadvantage (*Yiallourou* et al., *2022*). This is likely to be contingent on confounding factors that could not be controlled for in these analyses. Within each IRSD fifth, treated women may be less disadvantaged than untreated women, and this may have become more evident in the lower fifths. Also, fertility treatment is independently associated with some pregnancy complications, such as pre-eclampsia (*Watanabe* et al., *2014*), which in turn is associated with long-term adverse CVD outcomes (*Wu* et al., *2017*). Future research should include mediation analysis to determine how much of the effect is mediated by pregnancy complications, and how much is the independent effect of fertility treatment.

The finding of a reduced risk of CVD hospitalization among treated women who produced an average of fewer than five oocytes per cycle may represent the 'healthy patient' phenomenon previously reported in our study on mortality and in other work conducted on cohorts of women who received fertility treatments (*Venn* et al., *2001*). One of the major reasons for not proceeding to fertility treatment is financial difficulty (*Kulkarni* et al., *2014*). As such, women in our cohort who received treatment may have been wealthier and, therefore, healthier than those who did not, owing ro factors such as increased health literacy, greater access to health-influencing factors, including nutritious foods, and greater access to higher-quality health services, on behalf of their relative financial stability. This could explain their decreased propensity to be subsequently hospitalized for CVD. If this was the case, we would expect to find the same reduced risk in those women who averaged fewer than five oocytes, which we did not. It is possible that these women who have a stronger response to fertility drugs may have more acute complications, such as ovarian hyperstimulation, which may elevate their long-term CVD risk. Our finding that treated women who averaged fewer than five oocytes were at increased risk of CVD hospitalization, compared with treated women who averaged fewer than five oocytes, supports this notion. The higher incidence of CVD among women who average fewer than five oocytes may also relate to polycystic ovary sydrome, which is associated with larger oocyte numbers, being more prevalent among this group (*Weghofer* et al., *2007*). Differences in socioeconomic status between treated and untreated women may have confounded the comparison, especially as financial difficulties were among the most common reasons for not proceeding to treatment. The fact that unexplained infertility was more than twice as prevalent in untreated compared with treated women suggests that other factors, e.g. natural conception, age at registration, may also be implicated.

These results have meaningful clinical implications for the millions of women who will consider fertility treatments as technologies continue to advance. Counselling is an essential component of the clinical management of these women, each of whom will have an individual profile of CVD and risk factors that need to be considered. The findings from this study, and other similar analyses, which found no elevated risks of adverse long-term CVD outcomes in women who undergo fertility treatment, provide a degree of reassurance for women and their clinicians. Caution should be exercised in management of women who are at higher risk of treatment failure, given the previously reported adverse findings in this population group.

Strengths and limitations

The major strengths of this study include the large sample size of 27,262 women; the ability to stratify analysis by factors including area-based social disadvantage, stimulated cycles and oocytes per cycle; and the granular analysis of CVD subtypes at ICD-10 three-digit diagnosis level. One of the limitations was that information on lifestyle, e.g. diet, smoking, exercise, and certain biochemical factors, e.g. HbA1c, cholesterol, which are related to CVD risk, was unavailable so the analysis could not adjust for these potential confounders. Furthermore, it was not possible to determine whether some of the women who opted not to undergo fertility therapy at Monash IVF clinics were subsequently treated elsewhere, so some women may have been erroneously misclassified as 'untreated'. Also, the treated women were not followed up to determine whether they had successful pregnancies. As pregnancy itself contributes to CVD risk (*Smith* et al., *2019*), this may have confounded the results if women in the treated group were more likely to have conceived during our study period. Finally, as this study only included women from a specific IVF clinic in Australia, the results may not be generalizable to women receiving fertility treatment in other jurisdictions.

In conclusion, fertility treatment did not increase the risk of CVD hospitalization, overall or by subtype, in a cohort of 27,262 women. Our findings add to the growing body of evidence suggesting an absence of long-term CVD risk associated with fertility treatment. In conjunction with similar findings from other studies across the globe, these results provide reassurance that fertility treatment is a safe option for many women.

DATA AVAILABILITY

Data will be made available on request.

ACKNOWLEDGEMENTS

This study was funded by a Monash IVF Research and Education Foundation grant and by the Victorian Government Operational Infrastructure Support Program. Dunya Tomic was supported by an Australian Government Research Training Program Scholarship and Monash Graduate Excellence Scholarship. Dianna Magliano was supported by a National Health and Medical Research Council Senior Research Fellowship. Melinda Carrington received an endowed fellowship in the Cardiology Centre of Excellence from Filippo and Maria Casella.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103812.

REFERENCES

Australian Bureau of Statistics, 2022. Socio-Economic Indexes for Areas. Available from: https://wwwabs.gov.au/websitedbs/censushome.nsf/home/seifa (accessed 23 June 2022).

Ben-Yaakov, R.D., Kessous, R., Shoham-Vardi, I., Sergienko, R., Pariente, G., Sheiner, E., 2016. Fertility Treatments in Women Who Become Pregnant and Carried to Viability, and the Risk for Long-Term Maternal Cardiovascular Morbidity. Am J Perinatol 33, 1388–1393.

Dayan, N., Filion, K.B., Okano, M., Kilmartin, C., Reinblatt, S., Landry, T., Basso, O., Udell, J.A., 2017. Cardiovascular Risk Following Fertility Therapy: Systematic Review and Meta-Analysis. J Am Coll Cardiol 70, 1203–1213.

Eskew, A.M., Jungheim, E.S., 2017. A History of Developments to Improve in vitro Fertilization. Mo Med 114, 156–159.

Fraser, A., Nelson, S.M., Macdonald-Wallis, C., Cherry, L., Butler, E., Sattar, N., Lawlor, D.A., 2012. Associations of pregnancy complications with calculated cardiovascular disease risk and cardiovascular risk factors in middle age: the Avon Longitudinal Study of Parents and Children. Circulation 125, 1367–1380.

Hart, R., Doherty, D.A., 2015. The potential implications of a PCOS diagnosis on a woman's long-term health using data linkage. J Clin Endocrinol Metab 100, 911–919.

Kulkarni, G., Mohanty, N.C., Mohanty, I.R., Jadhav, P., Boricha, B.G., 2014. Survey of reasons for discontinuation from in vitro fertilization treatment among couples attending infertility clinic. J Hum Reprod Sci 7, 249–254.

Magnus, M.C., Fraser, A., Haberg, S.E., Rono, K., Romundstad, L.B., Bergh, C., Spangmose, A.L., Pinborg, A., Gissler, M., Wennerholm, U., Asvold, B.O., Lawlor, D.A., 2023. Opdahl S. Maternal risk of cardiovascular disease after use of assisted reproductive technologies. JAMA Cardiol 8, 837–845.

Patrizio, P., Albertini, D.F., Gleicher, N., Caplan, A., 2022. The changing world of IVF: the pros and cons of new business models offering assisted reproductive technologies. Journal of Assisted Reproduction and Genetics 39, 305–313.

Smith, G.N., Louis, J.M., Saade, G.R., 2019. Pregnancy and the Postpartum Period as an Opportunity for Cardiovascular Risk Identification and Management. Obstetrics & Gynecology 134.

Sneed, M.L., Uhler, M.L., Grotjan, H.E., Rapisarda, J.J., Lederer, K.J., Beltsos, A.N., 2008. Body mass index: impact on IVF success appears age-related. Human Reproduction 23, 1835–1839.

Tan, J., Taskin, O., Iews, M., Lee, A.J., Kan, A., Rowe, T., Bedaiwy, M.A., 2019. Atherosclerotic cardiovascular disease in women with endometriosis: a systematic review of risk factors and prospects for early surveillance. Reprod Biomed Online 39, 1007–1016.

Udell, J.A., Lu, H., Redelmeier, D.A., 2013. Long-term cardiovascular risk in women prescribed fertility therapy. J Am Coll Cardiol 62, 1704–1712.

University of New South Wales, 2020. Assisted reproductive technology in Australia and New Zealand 2018. Available from: https://npesuunsw.edu.au/surveillance/assisted-reproductive-technology-australia-and-new-zealand-2018 (accessed 12 June 2023).

Venn, A., Hemminki, E., Watson, L., Bruinsma, F., Healy, D., 2001. Mortality in a cohort of IVF patients. Hum Reprod 16, 2691–2696.

Watanabe, N., Fujiwara, T., Suzuki, T., Jwa, S.C., Taniguchi, K., Yamanobe, Y., Kozuka, K., Sago, H., 2014. Is in vitro fertilization associated with preeclampsia? A propensity score matched study. BMC Pregnancy Childbirth 14, 69.

Weghofer, A., Munne, S., Chen, S., Barad, D., Gleicher, N., 2007. Lack of association between polycystic ovary syndrome and embryonic aneuploidy. Fertil Steril 88, 900–905.

Westerlund, E., Brandt, L., Hovatta, O., Wallén, H., Ekbom, A., Henriksson, P., 2014. Incidence of hypertension, stroke, coronary heart disease, and diabetes in women who have delivered after in vitro fertilization: a population-based cohort study from Sweden. Fertil Steril 102, 1096–1102.

Wu, P., Haththotuwa, R., Kwok, C.S., Babu, A., Kotronias, R.A., Rushton, C., Zaman, A., Fryer, A.A., Kadam, U., Chew-Graham, C.A., Mamas, M.A., 2017. Preeclampsia and Future Cardiovascular Health. Circulation: Cardiovascular Quality and Outcomes 10: e003497.

Yiallourou, S.R., Magliano, D., Haregu, T.N., Carrington, M.J., Rolnik, D.L., Rombauts, L., Rodrigues, A., Ball, J., Bruinsma, F.J., Da Silva Costa, F., 2022. Long term all-cause and cardiovascular disease mortality among women who undergo fertility treatment. Med J Aust 217, 532–537.

Zhang, M., Li, J., Fu, X., Zhang, Y., Zhang, T., Wu, B., Han, X., Gao, S., 2022. Endometrial thickness is an independent risk factor of hypertensive disorders of pregnancy: a retrospective study of 13,458 patients in frozen-thawed embryo transfers. Reprod Biol Endocrinol 20, 93.

Received 11 October 2023; received in revised form 1 January 2024; accepted 3 January 2024.

RBMO

ELSEVIER

ARTICLE

Screening women in young adulthood for disabling dysmenorrhoea: a nationwide cross-sectional study from the CONSTANCES cohort

BIOGRAPHY

François Margueritte is an Obstetrician and Gynaecological Surgeon. His fields of interest are gynaecological surgery (oncological or benign conditions). As a PhD student his research focused on chronic pelvic pain in young adult women.

François Margueritte[1,2,3,*], Xavier Fritel[4,5], Annie Serfaty[1,6], Mireille Coeuret-Pellicer[7], Arnaud Fauconnier[1,3]

KEY MESSAGE

Assessing warning indicators such as intensity of dysmenorrhoea and associated pelvic pain symptoms and other correlates may help to detect, without medical knowledge, young women who may have disabling dysmenorrhoea. This might help to identify young adult women who need medical advice to prevent the long-term consequences of disabling dysmenorrhoea.

ABSTRACT

Research question: How do different warning indicators help to identify disabling dysmenorrhoea among women in young adulthood?

Design: A nationwide cross-sectional study of women aged 18–25 years from the CONSTANCES cohort was constructed. Disability was assessed with the Global Activity Limitation Indicator question 'For the past 6 months, have you been limited in routine activities?Yes, severely limited/Yes, limited/ No, not limited'. Dysmenorrhoea pain intensity and other chronic pelvic pain symptoms (dyspareunia and non-menstrual pain) were evaluated according to questions from a specific questionnaire. Probability of disability was estimated using a logistic prediction model according to dysmenorrhoea intensity, other indicators of pelvic pain symptoms and other obvious covariates. The results of the predictive model of disabling dysmenorrhoea were presented on a nomogram.

Results: Among 6377 women, the rate of disability was estimated at 7.5%. Increased intensity of dysmenorrhoea (odds ratio [OR] 1.08, 95% confidence interval [CI] 1.04–1.13), increased frequency of dyspareunia (from OR 1.69, 95% CI 1.33–2.14 up to OR 3.41, 95% CI 2.16–5.38) non-menstrual chronic pelvic pain (OR 1.75, 95% CI 1.40–2.19), body mass index over 25 kg/m^2 (OR 1.45, 95% CI 1.17–1.80) and non-use of the hormonal contraceptive pill (OR 1.29, 95% CI 1.05–1.59) were significantly

[1] Team RISCQ 'Clinical risk and security on women's health and perinatal health', Université Paris-Saclay, UVSQ, Montigny-le-Bretonneux, France
[2] Primary Care and Prevention Team, CESP, INSERM, Villejuif, France
[3] Department of Gynecology and Obstetrics, Intercommunal Hospital Center of Poissy-Saint-Germain-en-Laye, Poissy, France
[4] Department of Obstetrics and Gynaecology, La Miletrie University Hospital, Poitiers, France
[5] INSERM CIC 1402, Poitiers University, Poitiers, France
[6] Department of Medical Information, Territorial Hospital Group (GHT), Aisne-Nord/Haute-Somme, Saint Quentin Hospital, Aisne, France
[7] Population-Based Epidemiological Cohorts, UMS 11, Paris-Saclay University, Versailles St Quentin University, Université de Paris, INSERM, Villejuif, France

KEY WORDS
Dysmenorrhoea
Disability
Dyspareunia
Non-menstrual chronic pelvic pain
Young adulthood

*Corresponding author. E-mail address: fmargueritte@gmail.com (F. Margueritte). https://doi.org/10.1016/j.rbmo.2024.103861 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

associated with disability. According to the nomogram, a predicted probability of 15% or more could be chosen as a threshold. This represents almost 4.6% of young women in this sample being classified at risk of disabling dysmenorrhoea.

Conclusions: Dysmenorrhoea pain intensity and associated pelvic pain symptoms are warning indicators that can be measured to help screen young women who may suffer from disabling dysmenorrhoea.

INTRODUCTION

Dysmenorrhoea is known as an extremely common condition among women of childbearing age. Nearly 90% of women overall, and as many as 92% of women aged between 18 and 25 years, are affected by it (*Armour et al., 2020*; *Durand et al., 2021*; *Fauconnier et al., 2006*; *Jamieson and Steege, 1996*; *Margueritte et al., 2021*; *Pitts et al., 2008*; *Subasinghe et al., 2016*; *Zondervan et al., 1999*). The 18- to 25-year-old age group is characterized by the transition between adolescence and adulthood (*Sawyer et al., 2018*), a critical stage of development that could be disturbed by chronic pelvic pain (CPP) symptoms such as dysmenorrhoea.

However, only a third of women with dysmenorrhoea seek medical help (*Durand et al., 2021*; *Parker et al., 2010*). The rest of do not seek medical advice or support because they believe, or have been taught, that dysmenorrhoea is part of a woman's lot (*Chen et al., 2018*; *Wong, 2011*). Instead, these women often resort to self-medication or self-care therapies to relieve their symptoms (*Armour et al., 2019a*). In addition, the lack of training of teachers and education staff to discuss painful periods has not been helpful (*Armour et al., 2022*). Therefore, the issue of dysmenorrhoea cannot be reduced to simply a question of pain intensity: the impact on school performance and attendance, social isolation and loss of productivity should also be accounted for (*Armour et al., 2020*; *Parker et al., 2010*).

The population-based detection of women with severe dysmenorrhoea could be useful for several reasons: avoiding the individual and societal consequences of related limitations on activity and avoiding complications such as progression towards sensitization (*Aredo et al., 2017*; *Lamvu et al., 2021*; *Stratton et al., 2015*). For young women with dysmenorrhoea, treating the symptoms in order to avoid chronic pain and its chronic complications seems to be more important than finding the cause (*Iacovides et al., 2015*). However, when dysmenorrhoea is associated with other CPP symptoms such as dyspareunia and non-menstrual CPP (NMCPP), this could also suggest underlying pathologies such as endometriosis (*Becker et al., 2022*; *Chapron et al., 2003*; *Zondervan et al., 2020*).

Whether dysmenorrhoea is a physiological or pathological condition is debatable, but given the high prevalence of dysmenorrhoea, the real questions should be, how can it be measured and what are the consequences in terms of disability? To assess the measurement of dysmenorrhoea, warning indicators therefore need to be identified that may enable the classification of women who are or are not at risk of further complications (*Armour et al., 2019a*; *Durand et al., 2021*). These indicators have to be easy to evaluate (single-question symptoms, for example) and used by everyone, including people without medical knowledge.

The aim of the current study was to assess warning indicators of disabling dysmenorrhoea to allow the identification of a specific population of young women who could be proposed for specific medical counselling.

MATERIALS AND METHODS

Population setting

The study's population is based on the CONSTANCES cohort study (*Goldberg et al., 2017*). This cohort has involved more than 200,000 people affiliated to the French national health insurance system. Participants were enrolled in the CONSTANCES cohort study, whose recruitment period lasted from the end of 2012 until the end of 2021.

This cohort was designed for epidemiological purposes. Participants were followed every year with a questionnaire and every 5 years with a full medical examination (*Ruiz et al., 2016*; *Zins et al., 2015*). To reduce the selection bias and discrepancies between age groups, people invited to participate were randomly selected according to their national insurance number, gender and age. When they agreed to participate, they were asked to complete numerous questionnaires about lifestyle, medical history, history of employment and other aspects of their life (e.g. tobacco or alcohol use, hearing impairment, breathing and history of employment) and offered free medical examination at a healthcare centre.

Each woman who agreed to take part in the cohort study was asked to complete a questionnaire about specific women's topics, especially the assessment of CPP symptoms. This questionnaire was developed before enrolment and is based on questions specifically designed to assess endometriosis-related pain and CPP symptoms: dysmenorrhoea, dyspareunia and NMCPP (*Jamieson and Steege, 1996*; *Peveler et al., 1996*).

The current study included women between 18 and 25 years old. Women who did not answer the question concerning dysmenorrhoea and who had not had periods in the previous 3 months were excluded from the study, as were those who did not answer the question concerning limitation of daily activities (FIGURE 1). All the questions that were used for this study (disability and assessment of CPP symptoms) are available in Supplemental Material 1.

Measurement of dysmenorrhoea severity and other CPP symptoms

All the questions about CPP symptoms were based on specific questionnaires that had previously been used and validated by the CONSTANCES cohort study to assess pain-related symptoms for endometriosis or CPP. The list of symptoms included in the cohort's questionnaires came from a comprehensive review of the literature and was constructed according to the verbal answers of women suffering from endometriosis (*Chapron et al., 2005*).

Severity of dysmenorrhoea severity was evaluated with a numerical rating scale from 0 'No pain' to 10 'Worst pain ever' (*Chapron et al., 2003*). This scale is useful and is recommended to the assess intensity of dysmenorrhoea (*Ameade and Mohammed, 2016*; *Bourdel et al., 2015*).

FIGURE 1 Flow chart of the study population, to build a predictive model assessing disability based on chronic pelvic pain symptoms in young women included in the CONSTANCES cohort (*Zins et al., 2015*).

NMCPP was evaluated by a yes/no question about the presence of CPP not associated with periods or intercourse (*Chapron et al., 2005*).

Dyspareunia was only assessed for women who had had intercourse at least once. Those who had never had intercourse were asked to skip the question. Respondents were asked to rate the frequency of their dyspareunia as 'never or exceptionally', 'occasionally', 'frequently' or 'always' (*Fauconnier et al., 2006*). They could also answer 'don't want to answer'. Women who never had intercourse or who chose 'don't want to answer' were set apart as a specific group.

Measurement of disability

Disability was evaluated using the Global Activity Limitation Indicator (GALI), constructed to assess health-related activity limitation and disability in Europe (*Robine et al., 2003*). This single-item indicator has been the subject of numerous adaptations and is now considered to be a good indicator of activity limitation (*Jagger et al., 2010*). Furthermore, it has been translated and validated in French (*Berger et al., 2015*).

The question is as follows: 'For the past 6 months, have you been limited, i.e. have you experienced difficulties due to a health-related problem, in performing routine activities (at home, at work, leisure activities, etc.) in comparison with other people of your age?' It was considered that women had a disability (i.e. were limited in their daily activities) when they answered 'severely' or 'limited but not severely', but did not have one when they answered 'not limited' or 'slightly limited' another possible modality added to the GALI question asked within the Constances Cohort.

Assessment of confounding factors

Many factors, such as body mass index (BMI), age and type of contraception (especially hormonal contraception), may affect the intensity of dysmenorrhoea (*Armour et al., 2019b*; *Latthe, 2006*; *Margueritte et al., 2021*; *Pitts et al., 2008*; *Söderman et al., 2019*). However, the association of dysmenorrhoea with age among young women is debatable: some studies have reported a linear association between a decrease in dysmenorrhoea intensity and age, and others, no association at all (*Armour et al., 2019b*; *Margueritte et al., 2021*; *Pitts et al., 2008*). The association between BMI and dysmenorrhoea and disability is well known (*Casillas-Clot et al., 2021*; *Latthe, 2006*; *Le Strat et al., 2020*), as is the role of the type of contraception used (*Haynes et al., 2018*; *Mosher et al., 2018*; *Pitts et al., 2008*; *Söderman et al., 2019*).

Age and BMI were assessed with the questionnaire and a clinical examination before inclusion. The type of contraception was assessed using multiple choice questions in various categories according to their contraceptive effects: 'none, mechanical (condom, cup, sterilization), non-hormonal intrauterine device, hormonal intrauterine device, implant, hormonal contraceptive pills'. A priori confounding factors were also selected given their association with disability and CPP symptoms:

- healthcare refusal (*Baggio et al., 2017*; *Durand et al., 2021*; *Matin et al., 2021*);
- presence of children at home – 'yes or no' (*Horner-Johnson et al., 2016*; *Latthe, 2006*; *Righarts et al., 2018*; *Zhang et al., 2019*);
- history of diseases that may cause limitations of activity – 'yes or no': history of cancer, renal disease, psychiatric disorders (depression or suicide attempts), diabetes or osteoarticular diseases (*Dauphin and Eideliman, 2021*; *Latthe, 2006*);
- geographical origin – 'France, Europe, North Africa, Sub-Saharan Africa, Asia, other' (*Berger et al., 2015*; *Grace and Zondervan, 2004*; *Le Strat et al., 2020*);
- level of education according to the International Standard Classification of Education (*Casillas-Clot et al., 2021*; *Latthe, 2006*; *Pitts et al., 2008*; *Rubio Valverde et al., 2021*; *UNESCO Institute for Statistics, 2012*);
- monthly household income – '<450€, between 450 and <1000€, between 1000 and <1500€, between 1500 and <2100€, between 2100 and <2800€, between 2800 and 4200€, >4200€' (*Casillas-Clot et al., 2021*; *Da Luz et al., 2018*; *Latthe, 2006*);
- employed or unemployed – 'yes or no' (*Cabrero-García et al., 2020*;

Casillas-Clot et al., 2021; *Da Luz et al., 2018*; *Latthe, 2006*);
- living with a partner 'no partner, partner living separately, partners living together' (*Kapadi and Elander, 2020*; *Latthe, 2006*; *Tubeuf, 2008*).

All these variables were introduced into the model according to the best codification assessed by likelihood ratio tests.

Each woman in the CONSTANCES cohort has a follow-up, including an annual questionnaire about lifestyle and other diseases that may have been diagnosed since then. A diagnosis of endometriosis was one of the conditions, and the number of women who reported having been diagnosed with endometriosis at the last available follow-up (year 2020) was also taken into account.

Statistical analyses

Women who reported that they were not limited in their daily activities (i.e. those who answered 'slightly limited' or 'not limited') were compared with those who declared they were limited (i.e. those who answered 'limited' or 'severely limited'). The Student's t-test and chi-squared test were used for continuous and categorical variables, respectively. Logistic regression was then performed to assess the association of disability (dependent variable) and dysmenorrhoea with other CPP components (dyspareunia, NMCPP).

Dysmenorrhoea intensity was coded and tested in multiple situations to assess the fit with disability: continuous, splines with or without fractional polynomials, and by classes using likelihood ratio tests to choose the best codification (*Harrell, 2015*). All the variables studied in this study were tested to find the best codification to be introduced into the modelization (linear, categorical, binary). This was assessed by likelihood ratio tests.

Multiple modelling regression strategies were applied to assess the best model using different strategies (Supplemental Material 2). Limitation was the dependent variable. The independent variables were as follows:

- Model 1: only the three variables assessing CPP were included. Age, BMI and/or type of contraception were included in the model only if a statistical difference was found between the groups of women with and without limitation for each variable.
- Model 2: this was the same as model 1 with adjustment for all variables previously listed as confounding factors for limitation. If age or BMI or contraception use had not been included in model 1, it was introduced in this model.
- Model 3: this was the same as model 1 with the confounding factors (those added in model 2) implemented as a disease risk score. To keep the effects of all the confounding factors without introducing all of them into the logistic regression as independent variables, a disease risk score was calculated for each woman. This was based on a logistic regression between disability (dependent variable) and all the confounding factors (independent variables). An inverse probability weighting technique was used to obtain weights for each women in the sample (*Granger et al., 2019*; *Leyrat et al., 2019*).

The discrimination performance of each model was evaluated by calculating its area under the curve (AUC) (*Royston et al., 2009*). An internal validation procedure was performed with a cross-validation procedure made by groups of 20, to prevent over-adjustment for the predictive performance of the model (*Altman and Royston, 2000*; *Luque-Fernandez et al., 2019*).

A suitable logistic model was chosen to assess the association between disability and CPP symptoms, and Wald tests to assess the potential interaction between dysmenorrhoea and the use of contraception or the BMI. The same process was performed with dyspareunia. Potential linear trends for frequency of dyspareunia were also assessed after finding the best linear model to fit the study's purpose.

To present the results, a Kattan nomogram was constructed with the final logistic regression model based on beta-coefficients, and using methods that have been described elsewhere (*Diblasio and Kattan, 2003*; *Filleron et al., 2018*; *Iasonos et al., 2008*).

To solve the problem of missing values, and to avoid selection bias, logistic regression was performed using a multiple imputation with 10 datasets and multivariate imputation with chained equations with all the variables used for the modelization and according to their characteristics (linear with upper and lower bond, logit regression, multiple logit regression) (*Rubin, 1987*).

Once the assessment of the severity of dysmenorrhoea or other CPP symptoms associated with disability was complete, an acceptable threshold of probability of limitation was defined to classify the participants and detect a specific sample of young women at risk of limitation.

All the analyses were performed using STATA Statistical Software, Release 15.1 (StataCorp, USA) with the multiple imputation (mi) estimate command to estimate all the results using multiple imputations. A nomogram was obtained using Excel (Microsoft corporation, USA) and SAS software, version 9.4 (SAS Institute Inc., USA) using the PROC SGPLOT program (*Yang, 2013*).

Details of ethical approval

The CONSTANCES cohort project has obtained the authorization of the National Data Protection Authority (authorization no. 910486, dated 3 March 2011). CONSTANCES has been approved by the National Council for Statistical Information, the National Medical Council and the Institutional Review Board of the National Institute for Medical Research-INSERM (authorization no. 01–011, dated 18 February 2009). All participants signed a written informed consent form.

RESULTS

At time of the current study, 115,353 women had been included in the CONSTANCES cohort (*Zins et al., 2015*). The current cohort study included 6377 women between 18 and 25 years old who were analysed to assess disability associated with dysmenorrhoea (FIGURE 1).

Almost 7.5% of the young women were classified as being limited in their daily activities (i.e. they answered 'limited' or 'severely limited'). Comparisons between those who were limited and those who were not showed that the mean intensity of dysmenorrhoea was different between the two groups ($P < 0.001$), as was the distribution of NMCPP ($P < 0.001$) and dyspareunia ($P < 0.001$) (TABLE 1). Unlike BMI ($P < 0.001$), the age distribution was not statistically different ($P = 0.261$) between the groups. The distribution was

TABLE 1 BASELINE CHARACTERISTICS OF WOMEN ACCORDING TO LIMITATION (*N* = 6377)

Characteristics	Not limited (*n* = 5900)	Limited or severely limited (*n* = 477)	*P*-value
Chronic pelvic pain components			
Dysmenorrhoea intensity (*n* = 6055), missing *n* = 322	4.0 (±2.5)	4.9 (±2.7)	<0.001
Dyspareunia (*n* = 6377)			
Never or exceptionally (*n* = 2619)	2486 (42.1)	133 (27.9)	<0.001
Never intercourse or don't want to answer (*n* = 1057)	969 (16.4)	88 (18.4)	
Occasionally (*n* = 1900)	1732 (29.4)	168 (35.2)	
Frequently (*n* = 523)	468 (7.9)	55 (11.5)	
Always (*n* = 156)	128 (2.2)	28 (5.9)	
Missing (*n* = 122)	117 (2.0)	5 (1.0)	
Non-menstrual chronic pelvic pain (*n* = 6377)			
No (*n* = 4981)	4675 (79.2)	306 (64.2)	<0.001
Yes (*n* = 1093)	952 (16.1)	141 (29.6)	
Missing (*n* = 303)	273 (4.6)	30 (6.3)	
Gynaecological and medical characteristics			
Contraception (*n* = 6377)			
None (*n* = 1308)	1185 (20.1)	123 (25.8)	<0.001
Mechanical (condom, cup, sterilization) (*n* = 767)	710 (12.0)	57 (11.9)	
Non-hormonal intrauterine device (*n* = 329)	307 (5.2)	22 (4.6)	
Hormonal intrauterine device (*n* = 92)	80 (1.4)	12 (2.5)	
Implant (*n* = 176)	155 (2.6)	21 (4.4)	
Contraceptive pill (*n* = 3435)	3224 (54.6)	211 (44.2)	
Missing (*n* = 270)	239 (4.1)	31 (6.5)	
History of diseases causing disability (*n* = 6377)[a]			
No (*n* = 5847)	5459 (92.5)	388 (81.3)	<0.001
Yes (*n* = 530)	441 (7.5)	89 (18.7)	
Healthcare refusal (*n* = 6377)			
No (*n* = 5178)	4862 (82.4)	316 (66.2)	<0.001
Yes (*n* = 1025)	883 (15.0)	142 (29.8)	
Missing (*n* = 174)	155 (2.6)	19 (4.0)	
Demographic characteristics			
Age (years) (*n* = 6377)			
18 (*n* = 112)	103 (1.7)	9 (1.9)	0.261
19 (*n* = 770)	722 (12.2)	48 (10.1)	
20 (*n* = 855)	798 (13.5)	57 (11.9)	
21 (*n* = 760)	699 (11.8)	61 (12.8)	
22 (*n* = 768)	722 (12.2)	46 (9.6)	
23 (*n* = 878)	812 (13.8)	66 (13.8)	
24 (*n* = 1023)	938 (15.9)	85 (17.8)	
25 (*n* = 1211)	1106 (18.7)	105 (22.0)	
Body mass index (kg/m^2) (*n* = 6263), missing *n* = 114	22.7 (±4.1)	23.6 (±5.1)	<0.001
Geographical origin (*n* = 6377)			
France (*n* = 5904)	5480 (92.9)	424 (88.9)	<0.001
Europe (other than France) (*n* = 160)	145 (2.5)	15 (3.1)	
North Africa (*n* = 105)	91 (1.5)	14 (2.9)	

(*continued on next page*)

TABLE 1 (*Continued*)

Characteristics	Not limited (*n* = 5900)	Limited or severely limited (*n* = 477)	*P*-value
Sub-Saharan Africa (*n* = 69)	61 (1.0)	8 (1.7)	
Asia (*n* = 49)	43 (0.7)	6 (1.3)	
Other (*n* = 61)	53 (0.9)	8 (1.7)	
Missing (*n* = 29)	27 (0.5)	2 (0.4)	
Child at home (*n* = 6377)			
No (*n* = 5815)	5389 (91.3)	426 (89.3)	<0.001
Yes (*n* = 206)	175 (3.0)	31 (6.5)	
Missing (*n* = 356)	336 (5.7)	20 (4.2)	
Living with partner (*n* = 6377)			
No partner (*n* = 2970)	2756 (46.7)	214 (44.9)	0.182
Partner living separately (*n* = 1420)	1326 (22.5)	94 (19.7)	
Partners living together (*n* = 1953)	1787 (30.3)	166 (34.8)	
Missing (*n* = 34)	31 (0.5)	3 (0.6)	
Economic and educational settings			
Level of education (n=6377)[b]			
Level 0–2 (*n* = 235)	202 (3.4)	33 (6.9)	<0.001
Level 3–4 (*n* = 3116)	2842 (48.2)	274 (57.4)	
Level 5 or higher (*n* = 2978)	2811 (47.6)	167 (35.0)	
Missing (*n* = 48)	45 (0.8)	3 (0.6)	
Work (*n* = 6377)			
Unemployed (*n* = 2992)	2760 (46.8)	232 (48.6)	<0.001
Employed (*n* = 3265)	3030 (51.4)	235 (49.3)	
Missing (*n* = 120)	110 (1.9)	10 (2.1)	
Monthly household income (*n* = 6377)			
Income <450€ (*n* = 413)	380 (6.4)	33 (6.9)	0.007
450€ < income <1000€ (*n* = 876)	790 (13.4)	86 (18.0)	
1000€ < income <1500€ (*n* = 839)	768 (13.0)	71 (14.9)	
1500€ < income <2100€ (*n* = 793)	736 (12.5)	57 (11.9)	
2100€ < income <2800€ (*n* = 773)	716 (12.1)	57 (11.9)	
2800€ < income <4200€ (*n* = 847)	791 (13.4)	56 (11.7)	
Income >4200€ (*n* = 427)	414 (7.0)	13 (2.7)	
Do not know (*n* = 682)	632 (10.7)	50 (10.5)	
Do not wish to answer (*n* = 367)	336 (5.7)	31 (6.5)	
Missing (*n* = 360)	337 (5.7)	23 (4.8)	

Results are reported as *n* (%), or mean (±SD) for continuous variables.

P-values were calculated using a chi-squared or Student's t-test.

[a] Disease causing disability: history of cancer, renal disease, psychiatric disorders (depression or suicide attempt), diabetes or osteoarticular diseases.

[b] Level of education was defined according to the International Standard Classification of Education classification (*UNESCO Institute for Statistics, 2012*): level 0–2, up to secondary education; level 3–4, up to post-secondary non-tertiary education; level 5 or higher, short-cycle tertiary or higher than a Bachelor's degree.

not different between the two groups for women living with their partner, not living with their partner or not having a partner ($P = 0.182$). In contrast, all others variables (contraception [$P < 0.001$], history of disease causing disability [$P < 0.001$], healthcare refusal [$P < 0.001$], geographical origin [$P < 0.001$], having a child at home [$P < 0.001$], level of education [$P < <0.001$], work [$P < 0.001$] and monthly household income [$P = 0.007$]) presented a significant difference of distribution between the not limited and limited groups.

For the last follow-up questionnaire in 2020, out of 6333 women enrolled in the cohort and included in this study, only 58 reported having been diagnosed with endometriosis, which represents 0.92% of the sample of women answering the question in the follow-up questionnaires.

TABLE 2 PREDICTIVE MODEL OF DISABILITY ACCORDING TO DYSMENORRHOEA INTENSITY, PELVIC PAIN FACTORS AND CORRELATES WITHIN THE THREE MODELIZATION STRATEGIES

Variable	Model 1: multivariate			Model 2: multivariate adjusted for confounding factors[a]			Model 3: multivariate with disease risk score[b]		
	OR	95% CI	*P*-value	OR	95% CI	*P*-value	OR	CI 95% CI	*P*-value
Dysmenorrhoea	1.08	1.04–1.13	<0.001	1.07	1.03–1.12	0.001	1.08	1.03–1.13	0.001
Dyspareunia									
No	1 (ref			1 (ref)			1 (ref)		
No intercourse/do not wish to answer	1.42	1.06–1.89	0.018	1.69	1.23–2.32	0.001	1.48	1.08–2.04	0.015
Sometimes	1.69	1.33–2.14	<0.001	1.61	1.26–2.05	<0.001	1.66	1.27–2.18	<0.001
Often	1.92	1.37–2.70	<0.001	1.84	1.30–2.61	0.001	1.82	1.25–2.65	0.002
Always	3.41	2.16–5.38	<0.001	3.22	2.00–5.17	<0.001	3.78	2.18–6.54	<0.001
Non-menstrual cyclic pain									
No	1 (ref)			1 (ref)			1 (ref)		
Yes	1.75	1.40–2.19	<0.001	1.47	1.17–1.85	0.001	1.47	1.14–1.90	0.003
Body mass index >25 kg/m^2									
No	1 (ref)			1 (ref)			1 (ref)		
Yes	1.45	1.17–1.80	0.001	1.27	1.02–1.59	0.032	1.38	1.07–1.77	0.012
Contraceptive pill									
Yes	1 (ref)			1 (ref)			1 (ref)		
No	1.29	1.05–1.59	0.015	1.14	0.92–1.42	0.220	0.90	0.71–1.14	0.380
AUC	0.653	0.627–0.679		0.711	0.687–0.735		0.642	0.616–0.669	
AUC after cross-validation	0.643	0.617–0.670		0.688	0.663–0.713		0.629	0.603 - 0.656	

Model 1: without adjustment; model 2, confounding variables introduced into the modelization; model 3, confounding variables included as a disease risk score.
[a] Confounding factor: gynaecological and medical characteristics (history of diseases causing disability, healthcare refusal), demographic characteristics (age, geographical origin, children at home, living with partner), economic and educational characteristics (level of education, employment, monthly household income).
[b] Risk disease score: calculated after performing logistic regression between disability and confounding factor and obtaining probabilities of disability for each woman. According to the disability, a sampling weight using each probability was assigned for each subject and was used in model 3 with the multivariate logistic model.
AUC, area under the curve; ref, reference category.

For the multiple imputation, all the logistic regressions were performed as stated in 'Materials and Methods'. Age was not included in model 1. Dysmenorrhoea had a linear code according to the likelihood ratio tests. BMI and use of contraception were associated with limitation and were simplified into dichotomous variables (BMI over 25 kg/m^2 yes/no and use of contraceptive pill yes/no) according to likelihood ratio tests. The results of the logistic regressions are presented in TABLE 2.

In each model, the dysmenorrhoea score was correlated with disability: for model 1, the odds ratio was 1.08 (95% CI 1.04–1.13). For dyspareunia, women who did not have intercourse or did not wish to answer also had an increased risk of limitation: model 1 odds ratio 1.42 (95% CI 1.06–1.89). The findings showed that the risk of limitation rose with the frequency of dyspareunia, with a significant linear trend ($P < 0.001$) and an increasing odds ratio. A BMI over 25 kg/m^2 and not using a contraceptive pill were associated with a higher frequency of disability: model 1, odds ratios 1.45 (95% CI 1.17–1.80) and 1.29 (95% CI 1.05–1.59), respectively.

Model 1 was chosen to construct the nomogram for the following reasons: model 2 integrated too many variables to be able to be kept for the nomogram, and the AUC of model 1 and model 3 were not significantly different ($P = 0.061$). Model 1 was reliable and simpler to use than model 3 with weighting techniques.

The nomogram was created using the logistic equation derived from model 1 (FIGURE 2). It gives a probability of limitation according to the presence or absence of CPP components, use of a contraceptive pill or BMI over 25 kg/m^2. For instance, a woman with a BMI of 20 kg/m^2 and a dysmenorrhoea score of 4/10, who does not wish to answer the question about dyspareunia, has no NMCPP and is taking a contraceptive pill has a 5% probability of disability. Similarly, a woman with a dysmenorrhoea intensity of 7, NMCPP and frequent dyspareunia, a BMI of 23 kg/m^2 and contraception with a copper intrauterine device would have a 19% probability of disability. Finally, a woman with a BMI of 40 kg/m^2 and a dysmenorrhoea intensity of 10, who is always experiencing dyspareunia and has NMCPP, and who does not use contraception, would have a 41% probability of disability.

Accordingly, women who present a predicted probability of disability above 15% represent 4.6% of the current sample of women aged 18–25 years old. This threshold of 15% of a predicted probability of disability may be used as a threshold to define a group of women at risk of disabling dysmenorrhoea who may be offered dedicated medical care.

DISCUSSION

In this population of randomly selected young women from the CONSTANCES

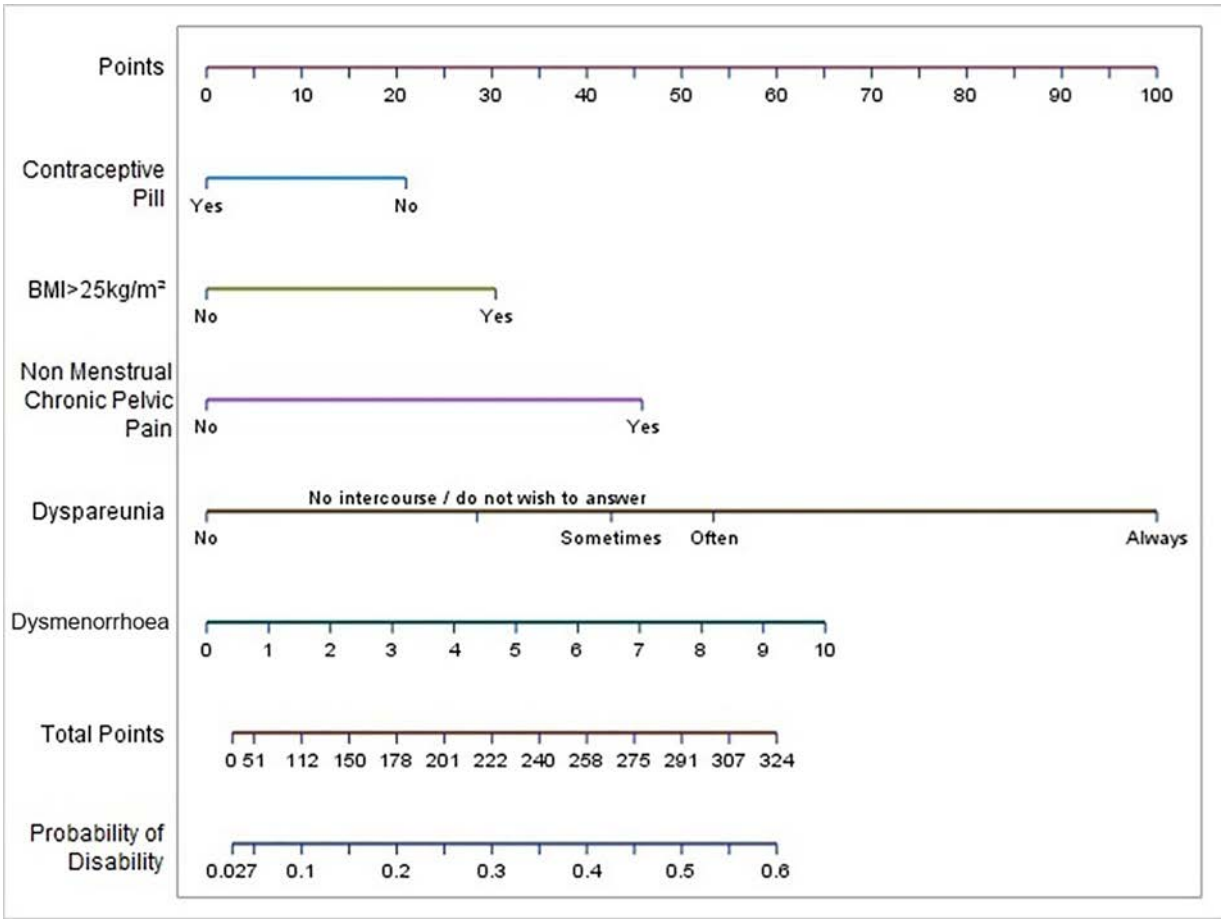


FIGURE 2 Nomogram estimating the probability of disability according to chronic pelvic pain symptoms and other correlates. Each individual's probability of limitation was estimated by plotting on each variable axis. A vertical line was drawn from that value to the top points scale to determine the number of points that were assigned by that variable value. Then the points from each variable value were summed. The sum of the total points scale was located and vertically projected onto the bottom axis, and then a personalized probability of limitation was obtained. BMI, body mass index.

cohort, it was found that 7.5% of women aged between 18 and 25 years had a disability associated with dysmenorrhoea and other CPP symptoms. Disability was strongly associated with increasing intensity of dysmenorrhoea, dyspareunia frequency and NCPP, and a high BMI (>25 kg/m^2). These criteria were combined to construct a nomogram in order to assess the probability of disability at a population level. A 15% threshold of probability was set to define a population at risk of presenting disability.

This represents twice the probability in the general population. This threshold could be used accordingly to subsequently define a population at risk of disability who could be offered medical support because it represents a small part of the population of young women. Nonetheless, to validate a strategy aimed at improving the screening of young women with disabling dysmenorrhoea, the nomogram must be prospectively evaluated using this threshold.

Many studies have shown that the prevalence of CPP symptoms is highest in young women (*Margueritte et al., 2021*; *Pitts et al., 2008*), and the current study has focused on the consequences of these symptoms. A highly reliable one-item tool (GALI) was used to detect disability in the general population (*Berger et al., 2015*) and to gain insights into disability associated with dysmenorrhoea. These results are consistent with findings in the Eurostat database, in France, using the same GALI indicators, with a proportion of 6.4–9% of women between 16 and 24 years old being limited or severely limited in their activities (*'Statistics | Eurostat,' n.d.*).

The CONSTANCES cohort has provided a robust population from which to construct a nomogram. Indeed, the women included in the current study were not selected to explore dysmenorrhoea or CPP symptoms; instead, they were randomly selected according to their social insurance number and received an invitation to participate that was sent by their national health insurance (*Zins et al., 2015*, *2010*). Before inclusion, the women did not know that they would have to answer questions about dysmenorrhoea, dyspareunia or menstruation, which provides a reliable sample of young women in France.

However, one selection bias could be that the women who were eager to participate did so because of the free medical examination and they possibly presented with more medical problems than the general population this study have tried to represent. Women who agreed to participate might be more interested in health topics and have a healthier lifestyle. However, in another study of the CONSTANCES cohort, using sampling weights (derived from a cohort of people who did not accept the invitation to participate) it was proved that taking these potential biases into account did not change the results about the prevalence of

CPP symptoms (*Margueritte et al., 2021*; *Santin et al., 2017*, *2014*).

The aim of this study was to develop a tool that could be used to screen young women according to their symptoms related to dysmenorrhoea. Nomograms are clinically useful tools that have been previously described for the purpose of screening for cancer (*Diblasio and Kattan, 2003*). In the current study, using a threshold of predicted probability of disability of 15% or above, only 4.6% of the study sample would be offered a personal medical examination relating to dysmenorrhoea and its correlates. A woman with a normal BMI and presenting two or more severe symptoms (dyspareunia always, a dysmenorrhoea intensity ≥8 or NMCPP) will have a probability of disability of more than 15% and would be offered medical advice. Women with a BMI of over 40 kg/m^2 and one severe symptom, especially severe dysmenorrhoea, without the use of a contraceptive pill, would also be proposed for medical support. This situation is typical of many women who suffer from primary dysmenorrhoea that is easily treated by contraceptive pills (*Armour et al., 2019a*; *Iacovides et al., 2015*). However, a woman who is presenting dysmenorrhoea of an intensity of 5 but has no other CPP, has a normal BMI and uses a contraceptive pill would not be offered a personal screening (as the probability of limitation is less than 15%).

This study also highlights the fact that, in young women, the problem is identifying disabling dysmenorrhoea rather than endometriosis, even when severe dysmenorrhoea is one of the symptoms. In this sample, only 0.92% of young women reported having been diagnosed with endometriosis at the most recent follow-up (for 6333 women), with no indication of how the diagnosis was made. Compared with the 90% of women with dysmenorrhoea or the 7.6% of women with disabling dysmenorrhoea, the question with a population of young women is how to detect disabling dysmenorrhoea in those who are not seeking medical help but are likely to need it. The aim of this strategy is to avoid chronic pain complications or long-term impairment of quality of life.

The originality of this study is that it focuses on the consequences of dysmenorrhoea, by studying disability associated with this symptom. As teachers say that they do not feel comfortable talking to their students about menstruation, it may not be appropriate for them to talk to pupils about disability related to dysmenorrhoea (*Duffy et al., 2013*). In addition, young women are not seeking medical help, so the extent of disability related to dysmenorrhoea is not really known.

The current study has used reliable tools and variables that make sense in clinical practice to create this nomogram. It could easily be used in a population of young women, by teachers or other non-medical staff. It does not require specific knowledge of disability or dysmenorrhoea. For example, a mobile application or QR code would ask just five questions, and if the result were above the threshold of 15%, it would offer women an appointment for medical support. After that, medical specialist staff would explain what part of the disability is due to dysmenorrhoea and its correlates, and how to treat them.

Parker and colleagues have also developed a dedicated screening tool for dysmenorrhoea (*Parker et al., 2022*). This is based on five binary questions about pelvic pain and feelings. Compared with the current authors' tool, that tool is just as useful, but it was developed from a sample of young women specifically selected for this purpose in a particular Australian area. A potential selection bias seems more obvious in the construction of this tool than in that of the current authors, even though the methodology and *statistics* are completely different.

An analysis of all the available data with multiple imputations allowed all the observations to be kept when constructing the modelization strategy (*Harrell, 2015*; *Rubin, 1987*). Without this statistical technique of multiple imputations, only 78% of the current sample would have been included in the logistic regression, leading to selection bias.

In this study's process of validation, use of a cross-validation has demonstrated that the results were reliable and unbiased (*Altman and Royston, 2000*; *Luque-Fernandez et al., 2019*). As the nomogram was constructed, the influence of potentially confounding factors that may disturb the association between disability and dysmenorrhoea was assessed. However, the introduction of confounding factors would have resulted in an over-complicated, cumbersome nomogram due to the many variables to present.

Instead, it was shown that the use of a disease risk score that estimates the weight of the many confounding factors for each participant did not change the validity of the initial model (model 1) that was chosen to build the nomogram.

A potential selection bias is that women of 18 years old were underrepresented in this study's sample. This is due to the difficulties of recruitment in this age group, despite the unequal inclusion probabilities used for drawing random samples from the population who are eligible for inclusion (*Zins et al., 2015*).

Finally, this model has a poor discriminatory power because the purpose was not to explain all the disability among young women. However, all the variables that were included and presented in the nomogram are clinically relevant and are strongly associated with disability. As the purpose was to identify women at risk of disabling dysmenorrhoea, the nomogram seems reliable in this case.

A level of 15% of the predicted probability of limitation corresponds to 4.6% of the population of young women in this study's sample. It therefore seems useful and feasible to carry out specific screening in the general population of young women using this specific tool, which could be used by anyone without specific knowledge. Other thresholds could have been discussed and need to be confirmed in a future specific prospective study.

CONCLUSIONS

This study has enabled the design of a nomogram that assesses the disability related to dysmenorrhoea by identifying several warning indicators. These could be used to identify women in need of medical advice to prevent the long-term consequences of disabling dysmenorrhoea, especially among young women. The performance and usefulness of this tool should be evaluated within different academic and non-academic structures to assess its utility and performance to detect women at risk of disabling dysmenorrhoea.

DATA AVAILABILITY

The authors do not have permission to share data.

ACKNOWLEDGEMENTS

The authors thank the 'Population-based Epidemiologic Cohorts Unit', UMS 011 Université de Paris, INSERM, Paris Saclay University, UVSQ, who designed and are in charge of the CONSTANCES project. They also thank the Caisse nationale de l'assurance maladie (CNAM) and the Centres d'examens de santé of the French Social Security scheme, who are collecting a large part of the data, as well as the Caisse nationale d'assurance vieillesse, ClinSearch, Asqualab and Eurocell, who are in charge of the data quality control. The authors also thank Anne Thiebaut and Elisabeth Delarocque-Astagneau for their advice while writing this article.

FUNDING

This project has received a grant from the Institute of Research of Women Health (www.irsf.fr) to get access to the data of the CONSTANCES cohort. None of the authors has been remunerated for this project. The CONSTANCES Project benefits from a grant from ANR (ANR-11-INBS-0002). CONSTANCES is also partly funded by MSD, AstraZeneca, Lundbeck and L'Oréal. None of these funding sources had any role in the design of the study, collection and analysis of data or decision to publish.

AUTHOR CONTRIBUTIONS

All the authors have made a substantial contribution to the preparation of the manuscript and have approved the final article.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103861.

REFERENCES

Altman, D.G., Royston, P., 2000. What do we mean by validating a prognostic model? Stat. Med. 19, 453–473. https://doi.org/10.1002/(sici)1097-0258(20000229)19:4<453::aid-sim350>3.0.co;2-5.

Ameade, E.K., Mohammed, B.S., 2016. Menstrual pain assessment: comparing verbal rating scale (VRS) with numerical rating scales (NRS) as pain measurement tools.

Aredo, J.V., Heyrana, K.J., Karp, B.I., Shah, J.P., Stratton, P., 2017. Relating Chronic Pelvic Pain and Endometriosis to Signs of Sensitization and Myofascial Pain and Dysfunction. Semin. Reprod. Med. 35, 88–97. https://doi.org/10.1055/s-0036-1597123.

Armour, M., Ferfolja, T., Curry, C., Hyman, M.S., Parry, K., Chalmers, K.J., Smith, C.A., MacMillan, F., Holmes, K., 2020. The Prevalence and Educational Impact of Pelvic and Menstrual Pain in Australia: A National Online Survey of 4202 Young Women Aged 13-25 Years. J. Pediatr. Adolesc. Gynecol. 33, 511–518. https://doi.org/10.1016/j.jpag.2020.06.007.

Armour, M., Parry, K., Al-Dabbas, M.A., Curry, C., Holmes, K., MacMillan, F., Ferfolja, T., Smith, C.A., 2019a. Self-care strategies and sources of knowledge on menstruation in 12,526 young women with dysmenorrhea: A systematic review and meta-analysis. PloS One 14, e0220103. https://doi.org/10.1371/journal.pone.0220103.

Armour, M., Parry, K., Curry, C., Ferfolja, T., Parker, M.A., Farooqi, T., MacMillan, F., Smith, C.A., Holmes, K., 2022. Evaluation of a web-based resource to improve menstrual health literacy and self-management in young women. J. Psychosom. Res. 162, 111038. https://doi.org/10.1016/j.jpsychores.2022.111038.

Armour, M., Parry, K., Manohar, N., Holmes, K., Ferfolja, T., Curry, C., MacMillan, F., Smith, C.A., 2019b. The Prevalence and Academic Impact of Dysmenorrhea in 21,573 Young Women: A Systematic Review and Meta-Analysis. J. Womens Health 28, 1161–1171. https://doi.org/10.1089/jwh.2018.7615 2002.

Baggio, S., Iglesias, K., Fernex, A., 2017. Healthcare renunciation among young adults in French higher education: A population-based study. Prev. Med. 99, 37–42. https://doi.org/10.1016/j.ypmed.2017.02.002.

Becker, C.M., Bokor, A., Heikinheimo, O., Horne, A., Jansen, F., Kiesel, L., King, K., Kvaskoff, M., Nap, A., Petersen, K., Saridogan, E., Tomassetti, C., van Hanegem, N., Vulliemoz, N., Vermeulen, N., ESHRE Endometriosis Guideline Group, 2022. ESHRE guideline: endometriosis. Hum. Reprod. Open hoac009. https://doi.org/10.1093/hropen/hoac009 2022.

Berger, N., Van Oyen, H., Cambois, E., Fouweather, T., Jagger, C., Nusselder, W., Robine, J.-M., 2015. Assessing the validity of the Global Activity Limitation Indicator in fourteen European countries. BMC Med. Res. Methodol. 15, 1. https://doi.org/10.1186/1471-2288-15-1.

Bourdel, N., Alves, J., Pickering, G., Ramilo, I., Roman, H., Canis, M., 2015. Systematic review of endometriosis pain assessment: how to choose a scale? Hum. Reprod. Update 21, 136–152. https://doi.org/10.1093/humupd/dmu046.

Cabrero-García, J., Juliá-Sanchis, R., Richart-Martínez, M., 2020. Association of the global activity limitation indicator with specific measures of disability in adults aged below 65. Eur. J. Public Health 30, 1225–1230. https://doi.org/10.1093/eurpub/ckaa066.

Casillas-Clot, J., Pereyra-Zamora, P., Nolasco, A., 2021. Determinants of Disability in Minority Populations in Spain: A Nationwide Study. Int. J. Environ. Res. Public. Health 18, 3537. https://doi.org/10.3390/ijerph18073537.

Chapron, C., Barakat, H., Fritel, X., Dubuisson, J.-B., Bréart, G., Fauconnier, A., 2005. Presurgical diagnosis of posterior deep infiltrating endometriosis based on a standardized questionnaire. Hum. Reprod. Oxf. Engl. 20, 507–513. https://doi.org/10.1093/humrep/deh627.

Chapron, C., Fauconnier, A., Dubuisson, J.-B., Barakat, H., Vieira, M., Bréart, G., 2003. Deep infiltrating endometriosis: relation between severity of dysmenorrhoea and extent of disease. Hum. Reprod. Oxf. Engl. 18, 760–766.

Chen, C.X., Shieh, C., Draucker, C.B., Carpenter, J.S., 2018. Reasons women do not seek health care for dysmenorrhea. J. Clin. Nurs. 27, e301–e308. https://doi.org/10.1111/jocn.13946.

Da Luz, R.A., de Deus, J.M., Conde, D.M., 2018. Quality of life and associated factors in Brazilian women with chronic pelvic pain. J. Pain Res. 11, 1367–1374. https://doi.org/10.2147/JPR.S168402.

Dauphin, L., Eideliman, S., 2021. Élargir les sources d'étude quantitative de la population handicapée : Que vaut l'indicateur « GALI » ? 30.

Diblasio, C.J., Kattan, M.W., 2003. Use of nomograms to predict the risk of disease recurrence after definitive local therapy for prostate cancer. Urology 62, 9–18. https://doi.org/10.1016/j.urology.2003.09.029.

Duffy, B., Fotinatos, N., Smith, A., Burke, J., 2013. Puberty, health and sexual education in Australian regional primary schools: Year 5 and 6 teacher perceptions. Sex Educ 13, 186–203.

Durand, H., Monahan, K., McGuire, B.E., 2021. Prevalence and Impact of Dysmenorrhea Among University Students in Ireland. Pain Med. Malden Mass 22, 2835–2845. https://doi.org/10.1093/pm/pnab122.

Fauconnier, A., Dubuisson, J.-B., Foulot, H., Deyrolles, C., Sarrot, F., Laveyssière, M.-N., Jansé-Marec, J., Bréart, G., 2006. Mobile uterine retroversion is associated with dyspareunia and dysmenorrhea in an unselected population of women. Eur. J. Obstet. Gynecol. Reprod. Biol. 127, 252–256. https://doi.org/10.1016/j.ejogrb.2005.11.026.

Filleron, T., Chaltiel, L., Jouve, E., Cabarrou, B., Gilhodes, J., Lusque, A., Mery, E., Dalenc, F., Martinez, A., 2018. Les nomogrammes dans la pratique clinique : méthodologie, limites et intérêts. Bull. Cancer (Paris) 105, 15–24. https://doi.org/10.1016/j.bulcan.2017.10.024.

Goldberg, M., Carton, M., Descatha, A., Leclerc, A., Roquelaure, Y., Santin, G., Zins, M., team, CONSTANCES, 2017. CONSTANCES: a general prospective population-based cohort for occupational and environmental epidemiology: cohort profile. Occup. Environ. Med. 74, 66–71. https://doi.org/10.1136/oemed-2016-103678.

Grace, V.M., Zondervan, K.T., 2004. Chronic pelvic pain in New Zealand: prevalence, pain severity, diagnoses and use of the health services. Aust. N. Z. J. Public Health 28, 369–375.

Granger, E., Sergeant, J.C., Lunt, M., 2019. Avoiding pitfalls when combining multiple imputation and propensity scores. Stat. Med. 38, 5120–5132. https://doi.org/10.1002/sim.8355.

Harrell, J., 2015. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, 2nd ed. Springer International Publishing: Imprint, Springer, Cham. https://doi.org/10.1007/978-3-319-19425-7. 2015. ed, Springer Series in Statistics.

Haynes, R.M., Boulet, S.L., Fox, M.H., Carroll, D.D., Courtney-Long, E., Warner, L., 2018. Contraceptive use at last intercourse among reproductive-aged women with disabilities: an analysis of population-based data from seven states. Contraception 97, 538–545. https://doi.org/10.1016/j.contraception.2017.12.008.

Horner-Johnson, W., Darney, B.G., Kulkarni-Rajasekhara, S., Quigley, B., Caughey, A.B., 2016. Pregnancy among US women: differences by presence, type, and complexity of disability. Am. J. Obstet. Gynecol. 214,. https://doi.org/10.1016/j.ajog.2015.10.929 529.e1-529.e9.

Iacovides, S., Avidon, I., Baker, F.C., 2015. What we know about primary dysmenorrhea today: a critical review. Hum. Reprod. 21, 762–778. https://doi.org/10.1093/humupd/dmv039 Update.

Iasonos, A., Schrag, D., Raj, G.V., Panageas, K.S., 2008. How to build and interpret a nomogram for cancer prognosis. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 26, 1364–1370. https://doi.org/10.1200/JCO.2007.12.9791.

Jagger, C., Gillies, C., Cambois, E., Van Oyen, H., Nusselder, W., Robine, J.-M., Team, EHLEIS, 2010. The Global Activity Limitation Index measured function and disability similarly across European countries. J. Clin. Epidemiol. 63, 892–899. https://doi.org/10.1016/j.jclinepi.2009.11.002.

Jamieson, D.J., Steege, J.F., 1996. The prevalence of dysmenorrhea, dyspareunia, pelvic pain, and irritable bowel syndrome in primary care practices. Obstet. Gynecol. 87, 55–58.

Kapadi, R., Elander, J., 2020. Pain coping, pain acceptance and analgesic use as predictors of health-related quality of life among women with primary dysmenorrhea. Eur. J. Obstet. Gynecol. Reprod. Biol. 246, 40–44. https://doi.org/10.1016/j.ejogrb.2019.12.032.

Lamvu, G., Carrillo, J., Ouyang, C., Rapkin, A., 2021. Chronic Pelvic Pain in Women: A Review. JAMA 325, 2381–2391. https://doi.org/10.1001/jama.2021.2631.

Latthe, P., 2006. Factors predisposing women to chronic pelvic pain: systematic review. BMJ 332, 749–755. https://doi.org/10.1136/bmj.38748.697465.55.

Le Strat, Y., Melchior, M., Gorwood, P., Tebeka, S., Dubertret, C., 2020. The role of comorbidity in the association of obesity with unemployment and disability. Ann. Epidemiol. 45, 61–68. https://doi.org/10.1016/j.annepidem.2020.03.004.

Leyrat, C., Seaman, S.R., White, I.R., Douglas, I., Smeeth, L., Kim, J., Resche-Rigon, M., Carpenter, J.R., Williamson, E.J., 2019. Propensity score analysis with partially observed covariates: How should multiple imputation be used? Stat. Methods Med. Res. 28, 3–19. https://doi.org/10.1177/0962280217713032.

Luque-Fernandez, M.A., Redondo-Sánchez, D., Maringe, C., 2019. cvauroc: Command to compute cross-validated area under the curve for ROC analysis after predictive modeling for binary outcomes. Stata J. Promot. Commun. Stat. Stata 19, 615–625. https://doi.org/10.1177/1536867x19874237.

Margueritte, F., Fritel, X., Zins, M., Goldberg, M., Panjo, H., Fauconnier, A., Ringa, V., 2021. The Underestimated Prevalence of Neglected Chronic Pelvic Pain in Women, a Nationwide Cross-Sectional Study in France. J. Clin. Med. 10, 2481. https://doi.org/10.3390/jcm10112481.

Matin, B.K., Williamson, H.J., Karyani, A.K., Rezaei, S., Soofi, M., Soltani, S., 2021. Barriers in access to healthcare for women with disabilities: a systematic review in qualitative studies. BMC Womens Health 21, 44. https://doi.org/10.1186/s12905-021-01189-5.

Mosher, W., Hughes, R.B., Bloom, T., Horton, L., Mojtabai, R., Alhusen, J.L., 2018. Contraceptive use by disability status: new national estimates from the National Survey of Family Growth. Contraception 97, 552–558. https://doi.org/10.1016/j.contraception.2018.03.031.

Parker, M.A., Kent, A.L., Sneddon, A., Wang, J., Shadbolt, B., 2022. The Menstrual Disorder of Teenagers (MDOT) Study No. 2: Period ImPact and Pain Assessment (PIPPA) Tool Validation in a Large Population-Based Cross-Sectional Study of Australian Teenagers. J. Pediatr. Adolesc. Gynecol. 35, 30–38. https://doi.org/10.1016/j.jpag.2021.06.003.

Parker, M.A., Sneddon, A.E., Arbon, P., 2010. The menstrual disorder of teenagers (MDOT) study: determining typical menstrual patterns and menstrual disturbance in a large population-based study of Australian teenagers. BJOG Int. J. Obstet. Gynaecol. 117, 185–192. https://doi.org/10.1111/j.1471-0528.2009.02407.x.

Peveler, R., Edwards, J., Daddow, J., Thomas, E., 1996. Psychosocial factors and chronic pelvic pain: a comparison of women with endometriosis and with unexplained pain. J. Psychosom. Res. 40, 305–315. https://doi.org/10.1016/0022-3999(95)00521-8.

Pitts, M.K., Ferris, J.A., Smith, A.M.A., Shelley, J.M., Richters, J., 2008. Prevalence and correlates of three types of pelvic pain in a nationally representative sample of Australian women. Med. J. Aust. 189, 138–143.

Righarts, A., Osborne, L., Connor, J., Gillett, W., 2018. The prevalence and potential determinants of dysmenorrhoea and other pelvic pain in women: a prospective study. BJOG Int. J. Obstet. Gynaecol. 125, 1532–1539. https://doi.org/10.1111/1471-0528.15247.

Robine, J.-M., Jagger, C., Group, Euro-REVES, 2003. Creating a coherent set of indicators to monitor health across Europe: the Euro-REVES 2 project. Eur. J. Public Health 13, 6–14. https://doi.org/10.1093/eurpub/13.suppl_1.6.

Royston, P., Moons, K.G.M., Altman, D.G., Vergouwe, Y., 2009. Prognosis and prognostic research: Developing a prognostic model. BMJ 338, b604. https://doi.org/10.1136/bmj.b604.

Multiple Imputation for Nonresponse in Surveys. In: Rubin, D.B. (Ed.), Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA. https://doi.org/10.1002/9780470316696.

Rubio Valverde, J.R., Mackenbach, J.P., Nusselder, W.J., 2021. Trends in inequalities in disability in Europe between 2002 and 2017. J. Epidemiol. Community Health 75, 712–720. https://doi.org/10.1136/jech-2020-216141.

Ruiz, F., Goldberg, M., Lemonnier, S., Ozguler, A., Boos, E., Brigand, A., Giraud, V., Perez, T., Roche, N., Zins, M., 2016. High quality standards for a large-scale prospective population-based observational cohort: Constances. BMC Public Health 16, 877. https://doi.org/10.1186/s12889-016-3439-5.

Santin, G., Bénézet, L., Geoffroy-Perez, B., Bouyer, J., Guéguen, A., 2017. A two-phase sampling survey for nonresponse and its paradata to correct nonresponse bias in a health surveillance survey. Rev. Epidemiol. Sante Publique 65, 71–79. https://doi.org/10.1016/j.respe.2016.10.059.

Santin, G., Geoffroy, B., Bénézet, L., Delézire, P., Chatelot, J., Sitta, R., Bouyer, J., Gueguen, A., Cohorts Group, SNIIR-AM, 2014. In an occupational health surveillance study, auxiliary data from administrative health and occupational databases effectively corrected for nonresponse. J. Clin. Epidemiol. 67, 722–730. https://doi.org/10.1016/j.jclinepi.2013.10.017.

Sawyer, S.M., Azzopardi, P.S., Wickremarathne, D., Patton, G.C., 2018. The age of adolescence. Lancet Child Adolesc. Health 2, 223–228. https://doi.org/10.1016/S2352-4642(18)30022-1.

Söderman, L., Edlund, M., Marions, L., 2019. Prevalence and impact of dysmenorrhea in Swedish adolescents. Acta Obstet. Gynecol. Scand. 98, 215–221. https://doi.org/10.1111/aogs.13480.

Statistics | Eurostat [WWW Document], n.d. URL https://ec.europa.eu/eurostat/databrowser/view/HLTH_SILC_12__custom_5687743/default/table?lang=en (accessed 4.5.23).

Stratton, P., Khachikyan, I., Sinaii, N., Ortiz, R., Shah, J., 2015. Association of chronic pelvic pain and endometriosis with signs of sensitization and myofascial pain. Obstet. Gynecol. 125, 719–728. https://doi.org/10.1097/AOG.0000000000000663.

Subasinghe, A.K., Happo, L., Jayasinghe, Y.L., Garland, S.M., Gorelik, A., Wark, J.D., 2016. Prevalence and severity of dysmenorrhoea, and management options reported by young Australian women. Aust. Fam. Physician 45, 829–834.

Tubeuf, S., 2008. Social heterogeneity in self-reported health status and measurement of inequalities in health 26.

UNESCO Institute for Statistics, 2012. International standard classification of education: ISCED 2011. Comp. Soc. Res. 30.

Wong, L.P., 2011. Attitudes towards dysmenorrhoea, impact and treatment seeking among adolescent girls: a rural school-based survey. Aust. J. Rural Health 19, 218–223. https://doi.org/10.1111/j.1440-1584.2011.01213.x.

Yang, D., 2013. Build prognostic nomograms for risk assessment using SAS. Presented at the Proceedings of SAS Global Forum. Citeseer.

Zhang, Y., McLain, A.C., Davis, B., McDermott, S., 2019. Fecundity and Infertility Among Women with Disabilities in the United States. J. Womens Health 28, 934–940. https://doi.org/10.1089/jwh.2018.7267 2002.

Zins, M., Bonenfant, S., Carton, M., Coeuret-Pellicer, M., Guéguen, A., Gourmelen, J., Nachtigal, M., Ozguler, A., Quesnot, A., Ribet, C., Rodrigues, G., Serrano, A., Sitta, R., Brigand, A., Henny, J., Goldberg, M., 2010. The CONSTANCES cohort: an open epidemiological laboratory. BMC Public Health 10, 479. https://doi.org/10.1186/1471-2458-10-479.

Zins, M., Goldberg, M., CONSTANCES team, 2015. The French CONSTANCES population-based cohort: design, inclusion and follow-up. Eur. J. Epidemiol.. https://doi.org/10.1007/s10654-015-0096-4.

Zondervan, K.T., Becker, C.M., Missmer, S.A., 2020. Endometriosis. N. Engl. J. Med. 382, 1244–1256. https://doi.org/10.1056/NEJMra1810764.

Zondervan, K.T., Yudkin, P.L., Vessey, M.P., Dawes, M.G., Barlow, D.H., Kennedy, S.H., 1999. Prevalence and incidence of chronic pelvic pain in primary care: evidence from a national general practice database. Br. J. Obstet. Gynaecol. 106, 1149–1155.

Received 30 September 2023; received in revised form 13 January 2024; accepted 30 January 2024.

RBMO

ELSEVIER

ARTICLE

First report on successful delivery after retransplantation of vitrified, rapid warmed ovarian tissue in Europe

BIOGRAPHY

Nicole Sänger is Professor/Medical Director of the Centre for Reproductive Medicine at the University Hospital Bonn, Germany. She is a board member of the German Society of Gynaecology and Obstetrics, FertiPROTEKT and the German Society of Endocrinology and Reproductive Medicine. Her current research activities focus on fertility preservation and cryopreservation.

Nicole Sänger*, Julia John, Rebekka Einenkel, Andreas Schallmoser

KEY MESSAGE

Our findings are consistent with those of Suzuki (2015) and Silber (2018), increasing published deliveries after ovarian tissue vitrification, warming and transplantation to five worldwide – in Japan (two), USA (two) and Europe (one). As other advantages of vitrification prevail, for example cost- and time-effectiveness, we propose vitrification/rapid warming over slow freezing/thawing.

ABSTRACT

Research question: Cryopreservation of ovarian tissue is one feasible option to preserve female fertility prior to cancer treatment. The slow freezing protocol represents the current standard approach, while vitrification has been suggested as a promising alternative. This paper reports the follow-up and first successful delivery after retransplantation of vitrified, rapid warmed ovarian tissue in Europe.

Design: After the patient received a diagnosis of breast cancer, ovarian tissue was removed laparoscopically and sent via overnight transportation to University Hospital Bonn for vitrification on site. The patient was treated with chemotherapy, leading to ovarian failure. After 2 years, retransplantation of the vitrified, rapid warmed tissue was conducted on site.

Results: Two months after grafting, the patient reported regular menstrual cycles. After 1 further month a clinical pregnancy occurred, which ended in a spontaneous abortion at the 8th week of pregnancy. Six months after grafting, another naturally conceived pregnancy was determined, resulting in the birth of a healthy boy 14 months after retransplantation of the ovarian tissue.

Conclusions: Complementing the successful deliveries reported by the groups of Suzuki (Japan) and Silber (USA) regarding vitrified tissue, the current results confirm the high potential of this cryopreservation method in a clinical routine setting as an alternative approach to the widespread slow freezing method.

Department of Gynecological Endocrinology and Reproductive Medicine, University Hospital of Bonn, Germany

*Corresponding authors. E-mail address: nicole.saenger@ukbonn.de (N. Sänger). https://doi.org/10.1016/j.rbmo.2024.103940 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS
Fertility protection
Follicle
Ovarian tissue
Ovary
Slow freezing
Vitrification

INTRODUCTION

Cryopreservation of ovarian tissue is an important option in female fertility preservation prior to anti-cancer therapy or other gonadotoxic treatment (*Jadoul et al., 2017*; *Kim et al., 2018*; *Wallace et al., 2016*). Since the first live birth reported by Donnez and colleagues in 2004, multiple groups have reported successful deliveries after the cryopreservation, thawing and transplantation of ovarian tissue (*Donnez et al., 2004*). This demonstrates the high potential of this method mainly based on the slow freezing protocol (*Donnez et al., 2004*; *Donnez et al., 2015*; *Donnez et al., 2017*; *Hoekman et al., 2019*; *Jensen et al., 2017*; *Meirow et al., 2005*; *Meirow et al., 2016*; *Rodriguez-Wallberg et al., 2016*, *2017*; *Van der Ven et al., 2016*). The slow freezing method is currently the standard approach and is regarded as well established (*Anderson et al., 2020*). Over 140 live births following tissue reimplantation have been reported (*Diaz et al., 2022*).

However, slow freezing is time- and cost-consuming. Following the success of oocyte vitrification, the current authors aimed to assess the potential of ovarian tissue vitrification. Ovarian tissue vitrification (*Fabbri et al., 2014*; *Keros et al., 2009*; *Schallmoser et al., 2022*; *Schallmoser, Einenkel et al., 2023*; *Shi et al., 2017*; *Xiao et al., 2010*) is a promising alternative, substantiated by four successful deliveries reported by the groups of Suzuki (*Suzuki et al., 2015*; Japan) and Silber (*Silber et al., 2018*; USA). It serves as a time- and cost-saving alternative to slow freezing (*Schallmoser et al., 2022*). Moreover, it results in similar or even improved follicular survival and less DNA fragmentation (*Chang et al., 2011*; *Shi et al., 2017*; *Xiao et al., 2013*).

The success of this method depends on the exact protocol and especially on the ultrafast cooling rate. Therefore, a cryodevice consisting of a metal grid was chosen. In contrast to other approaches this provides ultrafast cooling rates along with a comfortably usable contact surface for variable tissue sizes (including biopsy punches). The current group's previous in-vitro results confirmed that vitrification with this cryodevice is as reliable as slow freezing in terms of follicle count (*Schallmoser et al., 2022*), onset of apoptosis and release of angiogenic factors (*Schallmoser, Einenkel et al., 2023*). Here, the first birth after ovarian tissue vitrification using the metal grid cryodevice combined with rapid vertical vitrification is reported.

MATERIALS AND METHODS

Ethics

The study was approved by the ethics commission of University Hospital Bonn (approval code 007/09; 21 February, 2023). The patient gave written, informed consent.

Patient characteristics

A 32-year-old nulliparous patient diagnosed with breast cancer was referred to University Hospital Bonn. Pathological biopsy analysis revealed non-invasive breast cancer staged cT2 (3.3 cm), cN0, M0, no amplification of HER-2 (triple-negative) and G3. The anti-Müllerian hormone concentration was 0.97 μg/l (normal range 1–8 μg/l; TABLE 1). Prior to chemotherapy, half of the right ovary was removed laparoscopically and transferred to University Hospital Bonn via overnight (*Schallmoser et al., 2023*) transportation at 4–8°C for processing, vitrification and storage on site.

The patient's chemotherapeutic treatment included four cycles of epirubicin and cyclophosphamide prior to weekly (12 ×) Taxol treatment, combined with carboplatin (AUC5). Menstrual bleeding stopped and post-menopausal gonadotrophin concentrations as well as hot flushes occurred. Transplantation of ovarian tissue was conducted 2.5 years after the initial surgery. The peripheral blood hormone values indicated premature ovarian insufficiency as well (TABLE 1).

Processing of ovarian tissue prior to vitrification

Ovarian medulla was dismantled leaving a fine, thin layer of stroma cells, and the ovarian cortex was cut into pieces measuring 10 mm × 5 mm × 1 mm (FIGURE 1B). Processing of the ovarian cortex was conducted in Custodiol (Dr. Franz Köhler Chemie, Germany) at 4°C on a cooling plate under a laminar air flow bench to maintain high quality standards. To determine the follicular viability of the cortex tissue prior to and after cryopreservation, two pieces of 2 mm biopsies (pfm medical, Germany) were taken (FIGURE 1C).

Decontamination of liquid nitrogen

Decontamination (*Parmegiani et al., 2010*) of liquid nitrogen was performed by exposure to UV irradiation at 254 nm (OSRAM Germicidial Puritec, Berlin, Germany) at a short distance on a sterile working bench.

Vitrification of ovarian tissue

High-throughput vitrification of ovarian cortex tissue was conducted as previously described (*Schallmoser et al., 2022*). In brief, cortex was transferred to GMOPS+ (Vitrolife, Sweden) supplemented with 10% serum substitute supplement (SSS; Fujifilm Irvine Scientific, USA) and 10% ethylene glycol (Merck, Germany) for a period of 5 min. The tissue was then equilibrated in GMOPS+ (Vitrolife, Sweden) supplemented with 10% SSS (Fujifilm Irvine Scientific, USA) and 20% ethylene glycol (Merck, Germany) for another 5 min. The final step included equilibration of the cortex tissue in GMOPS+ (Vitrolife, Sweden) supplemented with 10% SSS (Fujifilm Irvine Scientific, Santa Ana, USA), 35% ethylene glycol (Merck, Germany), 5% polyvinylpyrrolidone (PVP; Merck, Germany) and 0.5 mol/l sucrose (Merck, Germany) for 6 min. All equilibration steps were conducted in 5 ml medium/well in a 6-well dish (Sarstedt, Germany) on a rocking shaker at room temperature in a laminar flow cabinet (Mars 1200; Cooper Surgical, USA).

TABLE 1 LABORATORY RESULTS AFTER CHEMOTHERAPEUTIC TREATMENT AND PRIOR TO TRANSPLANTATION

Parameter	AMH (μg/l)	LH (mIU/ml)	FSH (mIU/ml)	Oestradiol (pg/ml)
Prior to chemotherapy	0.97	–	–	–
After chemotherapy (10 months after OTC and 7 months prior to transplantation)	0.01	32.1	57.2	6
Prior to transplantation	–	24.9	63.4	<5

AMH, anti-Müllerian hormone; OTC, ovarian tissue cryopreservation.

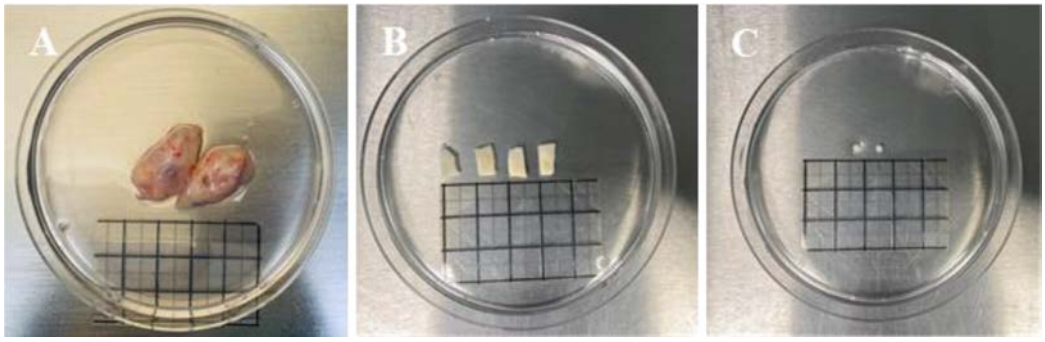


FIGURE 1 Representative images of the processing of the ovarian tissue: (A) removal of the medulla; (B) customized pieces; and (C) Two pieces of 2 mm biopsy punches for the determination of follicular viability before and after cryopreservation. Scale bar = 1 cm units.

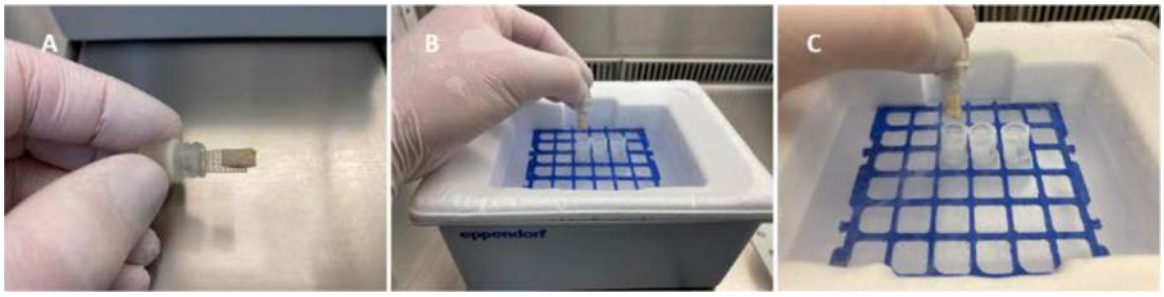


FIGURE 2 High-throughput vitrification of ovarian tissue. Processed ovarian cortex tissue on a metal grid (cryodevice) ready for vitrification (A, B, C). The grid-based Styrofoam (IsoSafe Box; Eppendorf, Germany) device enables a high sample throughput.

Within 60 s, cortex strips were loaded onto the cryodevice consisting of a sterile metal grid and submerged vertically into liquid nitrogen in prefilled cryovials arranged on a loading grid (FIG. 2). The samples were stored in the gas phase of liquid nitrogen (−160°C).

Rapid warming of ovarian tissue

The frozen samples were placed in a cryo-Dewar vessel containing liquid nitrogen. Keeping the lower part of the cryovials submerged in the liquid nitrogen, the caps were opened cautiously and the carriers with ovarian cortex tissue were submerged rapidly in a high volume (35 ml) of a 0.8 mol/l sucrose (Merck, Germany) solution consisting of GMOPS+ (Vitrolife, Sweden) and 10% SSS (Fujifilm Irvine Scientific, USA), at 37.2°C for 1 min. The tissue was then transferred to a 0.4 mol/l sucrose solution (Merck, Germany) consisting of GMOPS+ (Vitrolife, Sweden) and 10% SSS (Fujifilm Irvine Scientific, USA) at room temperature for 3 min prior to two washing steps (5 min each) in GMOPS+ (Vitrolife, Sweden) supplemented with 10% SSS (Fujifilm Irvine Scientific, USA), resulting in a thawing procedure lasting for 14 min.

Viability measurement

Two 2 mm ovarian cortex tissue punch pieces were digested for 90 min at 37.2°C in a solution consisting of 2 μmol/l calcein AM and 1 mg/ml collagenase type 1A (*Schallmoser, Einenkel et al., 2022*). The quantity of viable follicles was analysed with fluorescence microscopy (Ti2; Nikon, Germany) prior to cryopreservation and after thawing for quality control measurement as indicated in FIGURE 3.

Transplantation of ovarian tissue

Transplantation was performed 2 years and 5 months after explantation. Four pieces of ovarian cortex tissue were rapidly warmed and transferred into a peritoneal pocket, as indicated in FIGURE 4. A laparoscopy was performed for this purpose. The patency of the Fallopian tubes was checked by chromoperturbation.

The parietal peritoneum of the right ovarian fossa was incised over a length of

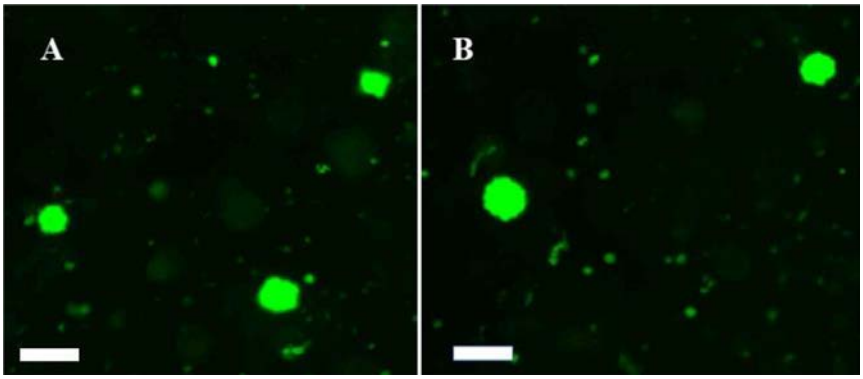


FIGURE 3 Determination of follicular viability. Green fluorescence (495 nm) is exhibited by granulosa cells surrounding the oocyte, indicating follicular viability, compared with the differently sized single cells from the ovarian stroma. Representative images of follicles prior to vitrification (A, nine viable follicles identified) and after rapid warming (B, eight viable follicles observed). Scale bar = 100 μm.

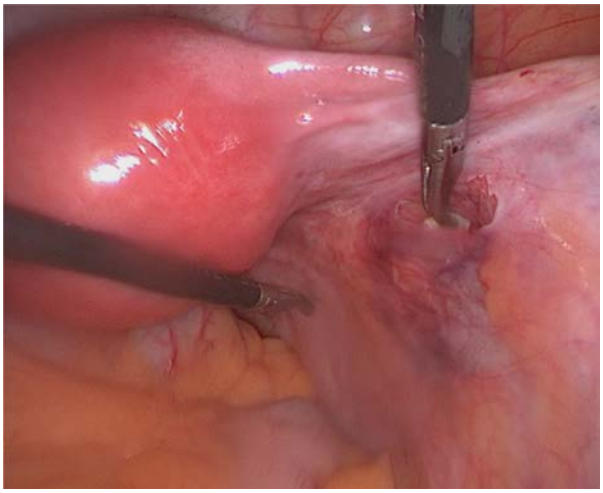


FIGURE 4 Transplantation of ovarian tissue into a peritoneal pocket close to the Fallopian tube. The ovarian cortical fragments are positioned with the cortical surface facing the abdominal cavity.

approximately 1.5 cm. A subperitoneal pocket was created close to the Fallopian tube by blunt dissection, with minimal bleeding, and the tissue pieces were implanted into the pocket with the cortical surface facing the abdominal cavity. The pocket was closed with a single suture.

RESULTS

Vitrification

Four ovarian cortex slices were vitrified, stored at −160°C in the gas phase of liquid nitrogen and rapidly warmed prior to transplantation.

Viability of ovarian tissue

Analysis of follicular viability with two pieces of 2 mm cortex punches gave nine viable follicles prior to cryopreservation. After the rapid warming of two 2 mm punch pieces, eight viable follicles were determined via fluorescence microscopy, resulting in a thawing survival rate of 88%. Non-viable follicles were not determined.

Outcome

Two months after transplantation of the ovarian tissue, regular menstrual cycles were reported by the patient, who conceived naturally, resulting in a clinical pregnancy that ended in an abortion at the 8th week of pregnancy (a fetal heartbeat could be detected sonographically). Six months after grafting, another naturally conceived pregnancy occurred. Because an antiphospholipid syndrome was suspected, therapy with heparin and aspirin was initiated. First- and second-trimester screening showed a healthy fetus. The patient gave birth via Caesarean section at term to a healthy boy 3601 g in weigh and 54 cm in length.

DISCUSSION

Vitrification revolutionized the cryopreservation of oocytes, with high survival rates. Although the results for human ovarian tissue are promising, only a few births have so far been reported after ovarian tissue vitrification and rapid warming.

Cryopreservation of ovarian tissue is an important option in terms of fertility preservation. Although 83% of the patients regain endocrine activity after reimplantation, only 20% give birth according to the data from the FertiPROTEKT network (*Barnitzky et al., 2023*). The stress of freezing and thawing as well as the undersupply after reimplantation, together with the reperfusion damage, impact follicular survival and lead to a significant decrease in the number of the follicles, limiting the success of the therapy option (*Cacciottola et al., 2021*). Vitrification is thought to preserve the follicles with a similar or even greater success compared with slow freezing approaches. Here, a patient with a relatively low follicular count regained ovarian function after only 2 months, evincing the success of the vitrification of human ovarian tissue.

There are varying techniques to assess the follicular count. In terms of the assessment of viable follicles by either neutral red staining or calcein staining coupled with tissue digestion, an average viable follicular count of 3.7–6.9 follicles/mm^2 in fresh tissue or 3.3–5.1 follicles/mm^2 in frozen-thawed tissue has been reported (*Cheng et al., 2021*; *Kristensen et al., 2018*). The current patient had a follicular density of 1.4 follicles/mm^2 before vitrification and 1.3 follicles/mm^2 after rapid warming.

Although the follicle count of the assessed tissue pieces were below average, the patient regained hormonal function quickly after transplantation. Despite the fact that the patient developed premature ovarian insufficiency with an FSH concentration of over 25 mIU/ml, amenorrhoea and a low oestradiol concentration, it only took 2 months for her ovarian function to recover after ovarian tissue reimplantation. This duration was below the reported average time to ovarian function recovery of 3.6–5.2 months (*Colmorn et al., 2022*; *Rozen et al., 2021*; *Shapira et al., 2020*). Moreover, the patient conceived naturally without the need for ovarian stimulation, oocyte retrieval or IVF/intracytoplasmic sperm injection. This report is, to the current authors' knowledge, the first live birth after vitrification, rapid warming and transplantation of ovarian tissue in Europe.

In the authors' opinion, the advantages of vitrification outweigh the disadvantages. In contrast to vitrification, traditional freezing approaches are associated with ice crystal formation, (*Amorim et al., 2011*; *Fahy et al., 1986*; *Liebermann et al., 2002*), potentially adversely affecting the tissue integrity and follicular viability. The often-discussed exposure to cryoprotectants, which might be detrimental during vitrification, clearly depends on the chosen protocol. The meta-analysis of Shi and colleagues describes the basic characteristics and outcomes of studies conducting ovarian tissue vitrification based on a huge variety of combined or singly used cryoprotectants such as ethylene glycol, dimethylsulfoxide (DMSO) and propanediol combined with metal-based, plastic-based and other carrier systems, which makes a direct comparison difficult (*Shi et al., 2017*).

Focusing on the reports with confirmed live births, the protocol used by Suzuki and co-workers is built upon increasing concentrations of ethylene glycol prior to loading on a metal carrier (*Suzuki et al., 2015*). The tissue is initially thawed with high concentrations of sucrose. With a certain degree of similarity, the procedure

performed by Silber and collaborators is based on rising concentrations of ethylene glycol and DMSO prior to loading onto thin metal strips (*Silber et al., 2018*). Thawing is conducted in high volumes of highly concentrated sucrose, followed by applying decreasing sucrose concentrations and final washing steps.

Consistent with these two reports, the results of the current group (*Schallmoser et al., 2023*; *Schallmoser, Einenkel et al., 2023*) show that ethylene glycol-based vitrification, combined with a metal-based loading system, with rapid warming in an initial step involving high concentrations of sucrose, seems to be a successful approach, confirming the high potential of ovarian tissue vitrification and thawing/rapid warming. This resembles to a certain extent the ethylene glycol-based vitrification of oocytes and embryos, which are thawed with decreasing concentrations of sucrose as well.

In rodents, (*Choi et al., 2008*; *Choi et al., 2013*) baboons (*Amorim et al., 2019*) and humans (*Xiao et al., 2013*; *Rahimi et al., 2009*; *Oktem et al., 2011*), partially conflicting results have been observed when comparing slow freezing and vitrification. This may be potentially caused by the use of different carrier systems and varying cryoprotectant and thawing agents.

The set of studies that have been performed to assess the feasibility of vitrification and compare it with slow freezing are heterogenous but favour vitrification, as shown in meta-analyses (*Behl et al., 2023*; *Shi et al., 2017*; *Zhou et al., 2016*). The meta-analyses included aspects such as the preservation of follicles, DNA double-strand breaks and the preservation of stromal cells, and found either no difference with or advantages of vitrification (*Behl et al., 2023*; *Shi et al., 2017*; *Zhou et al., 2016*).

Several aspects might explain the heterogeneity, one of which is the lack of a standardized vitrification protocol. Differences in the success of vitrification depend on multiple factors, which vary massively between these publications. These factors include the chosen cryoprotectants as well as the composition of the OTC media, the size of the ovarian fragments, the cryopreservation device (plastic or metal, shape and processing) and the speed of cooling. Furthermore, a standardized thawing protocol does not exist. Last but not least, whether differences can be detected or overlooked is a question of assessment. For example, the process of apoptosis can take up to 48 h and might not be detected when analysed earlier, (*Schallmoser et al., 2022*) especially since the follicles still seem morphologically intact (*Rimon et al., 2005*) in early apoptosis.

Moreover, vitrification is cost-effective in terms of cryopreservation time, in contrast to the slow freezing procedure that lasts several hours, increasing the chance of technical errors of the freezer unit, software and computer dysfunctions that can occur as a result of power failure or other deviations.

Finally, slow-freezing equipment such as computers, software and freezer units with attached nitrogen supply tanks, is associated with high acquisition costs and permanent maintenance fees, while the vitrification approach requires a minimum of equipment and maintenance – strongly resembling the vitrification process for oocytes and embryos. In contrast to slow freezing, however, vitrification requires experienced and well-trained staff.

In contrast to the vitrification of oocytes and embryos, very few commercially available Food and Drug Administration/Conformité Européenne (FDA/CE)-marked media and suitable carrier systems are available for the vitrification and thawing of ovarian tissue, leaving a commercial gap potentially hampering clinical routine. In IVF, regulatory frameworks recommend the use of FDA/CE-certified media (*Parmegiani et al., 2014*; *Parmegiani et al., 2018*). The current authors therefore recommend enhancing the development of optimized, commercially available, ready-to-use kits for ovarian tissue vitrification and thawing in order to take acquire experimental certification for this technique.

Finally these authors also subscribe, in the main, to the reasoning of Suzuki and co-workers (*Suzuki et al., 2015*) and Silber and collaborators (*Silber et al., 2018*) that vitrification of ovarian tissue is a favourable alternative to the traditional slow freezing approach.

DATA AVAILABILITY

Data will be made available on request.

REFERENCES

Amorim, C.A., Curaba, M., Langendonckt, A.V., Dolmans, M.M., Donnez, J., 2011. Vitrification as an alternative means of cryopreserving ovarian tissue. Reproductive BioMedicine Online 23, 160–186.

Amorim, C.A., Donnez, J., Dehoux, J.P., Scalercio, S.R., Squifflet, J., Dolmans, M.M. Long-term follow-up of vitrified and autografted baboon (Papio anubis) ovarian tissue. Hum Reprod. 2019;34(2):323-334. doi: 10.1093/humrep/dey355. PMID: 30551136.

Anderson, R.A., Amant, F., Braat, D., D'Angelo, A., Chuva de Sousa Lopes, S.M., Demeestere, I., Dwek, S., Frith, L., Lambertini, M., Maslin, C., Moura-Ramos, M., Nogueira, D., Rodriguez-Wallberg, K., Vermeulen, N., 2020. The ESHRE Guideline Group on Female Fertility Preservation, ESHRE guideline: female fertility preservation. Hum Reprod Open 2020 (4).

Barnitzky, S., Blumenauer, V., Czeromin, U., Fehr, D., Grewe, C., Krüssel, J.S., Kupka, M.S., Tandler-Schneider, A., Tauchert, S.D.I., 2023. R-Annual 2022. Journal of Reproductive Medicine and Endcrinology (5), 189–246.

Behl, S., Joshi, VB., Larson, NB., Young, MC., Bilal, M., Walker, DL., Khan, Z., Granberg, CF., Chattha, A., Zhao, Y., 2023. Vitrification versus slow freezing of human ovarian tissue: a systematic review and meta-analysis of histological outcomes. J Assist Reprod Genet 40 (3), 455–464. https://doi.org/10.1007/s10815-022-02692-w Epub 2022 Dec 21. PMID: 36542310; PMCID: PMC10033773.

Cacciottola, L., Donnez, J., Dolmans, M.M., 2021. Ovarian tissue damage after grafting: systematic review of strategies to improve follicle outcomes. Reprod Biomed Online 43 (3), 351–369. https://doi.org/10.1016/j.rbmo.2021.06.019 Epub 2021 Jun 26. PMID: 34384692.

Chang, H.J., Moon, J.H., Lee, J.R., Jee, B.C., Suh, C.S., Kim, S.H., 2011. Optimal condition of vitrification method for cryopreservation of human ovarian cortical tissues. J Obstet Gynaecol Res 37 (8), 1092–1101. https://doi.org/10.1111/j.1447-0756.2010.01496.x Epub 2011 Apr 19. PMID: 21501331.

Cheng, J., Ruan, X., Zhou, Q., Li, Y., Du, J., Jin, F., Gu, M., Mueck, A.O., 2021. Long-time low-temperature transportation of human ovarian tissue before cryopreservation. Reprod Biomed Online 43 (2), 172–183. https://doi.org/10.1016/j.rbmo.2021.05.006 Epub 2021 May 19. PMID: 34183267.

Choi, W.J., Lee, S.A., Lee, J.H., Han, J., Agarwal, A., Paik, W.Y., 2008. Comparison of angiopoietin levels in vitrified and slow-freezing mouse ovarian tissues. Fertil. Steril. 90, S280.

Choi, W.J., Lee, J.H., Park, M.H., Choi, I.Y., Park, J.K., Shin, J.K., Lee, S.A., Paik, W.Y., Lee, J.H., 2013. Influence of the vitrification solution on the angiogenic factors in vitrificated mouse ovarian tissue. Obstet Gynecol Sci 56 (6), 382–388. https://doi.org/10.5468/ogs.2013.56.6.382 NovEpub 2013 Nov 15. PMID: 24396817; PMCID: PMC3859015.

Colmorn, LB, Pedersen, AT, Larsen, EC, Hansen, AS, Rosendahl, M, Andersen, CY, Kristensen, SG, Macklon, KT., 2022. Reproductive and Endocrine Outcomes in a Cohort of Danish Women following Auto-Transplantation of Frozen/Thawed Ovarian

Tissue from a Single Center. Cancers (Basel) 14 (23), 5873. https://doi.org/10.3390/cancers14235873 PMID: 36497354; PMCID: PMC9740843.

Donnez, J., Dolmans, M.M., 2017. Fertility Preservation in Women. N Engl J Med 377, 1657–1665.

Diaz, A.A., Kubo, H., Handa, N., Hanna, M., Laronda, M.M., 2022. A Systematic Review of Ovarian Tissue Transplantation Outcomes by Ovarian Tissue Processing Size for Cryopreservation. Front Endocrinol (Lausanne) 13, 918899. https://doi.org/10.3389/fendo.2022.918899 PMID: 35774145; PMCID: PMC9239173.

Donnez, J., Dolmans, M.M., Demylle, D., Jadoul, P., Pirard, C., Squifflet, J., Martinez-Madrid, B., Van Langendonckt, A., 2004. Livebirth after orthotopic transplantation of cryopreserved ovarian tissue. Lancet 364, 1405–1410.

Donnez, J., Dolmans, M.M., Pellicer, A., Diaz-Garcia, C., Ernst, E., Macklon, K.T., Andersen, C.Y., 2015. Fertility preservation for age-related fertility decline. Lancet 385, 506–507.

Fabbri, R., Vicenti, R., Macciocca, M., Pasquinelli, G., Paradisi, R., Battaglia, C., Martino, N.A., Venturoli, S., 2014. Good preservation of stromal cells and no apoptosis in human ovarian tissue after vitrification. Biomed. Res. Int. 2014, 673537.

Fahy, G.M., 1986. Vitrification: a new approach to organ cryopreservation. In: Meryman, H.T. (Ed.), Transplantation: Approaches to Graft Rejection. Alan R. Liss, New York, USA, pp. 305–335.

Hoekman, E.J., Louwe, L.A., Rooijers, M., van der Westerlaken, L.A.J., Klijn, N.F., Pilgram, G.S.K., de Kroon, C.D., Hilders, C.G.J.M., 2019. Ovarian tissue cryopreservation: Low usage rates and high live-birth rate after transplantation. Acta Obstet Gynecol Scand 00, 1–9.

Jadoul, P., Guilmain, A., Squifflet, J., Luyckx, M., Votino, R., Wyns, C., Dolmans, M.M., 2017. Efficacy of ovarian tissue cryopreservation for fertility preservation: lessons learned from 545 cases. Human Reproduction 32 (5), 1046–1054 VolumeIssueMayPages.

Keros, V., Xella, S., Hultenby, K., Pettersson, K., Sheikhi, M., Volpe, A., Hreinsson, J., Hovatta, O., 2009. Vitrification versus controlled-rate freezing in cryopreservation of human ovarian tissue. Hum. Reprod. 24, 1670–1683.

Kim, H., Kim, H., Ku, S.Y., 2018. Fertility preservation in pediatric and young adult female cancer patients. Ann Pediatr Endocrinol Metab 23 (2), 70–74. https://doi.org/10.6065/apem.2018.23.2.70 PMID: 29969877; PMCID: PMC6057020.

Kristensen, SG, Liu, Q, Mamsen, LS, Greve, T, Pors, SE, Bjørn, AB, Ernst, E, Macklon, KT, Andersen, CY., 2018. A simple method to quantify follicle survival in cryopreserved human ovarian tissue. Hum Reprod 33 (12), 2276–2284. https://doi.org/10.1093/humrep/dey318 PMID: 30358835.

Liebermann, J., Nawroth, F., Isachenko, V., Isachenko, E., Rahimi, G., Tucker, M.J., 2002. Potential importance of vitrification in reproductive medicine. Biol. Reprod. 67, 1671–1680.

Meirow, D., Levron, J., Eldar-Geva, T., Hardan, I., Fridman, E., Zalel, Y., Schiff, E., Dor, J., 2005. Pregnancy after transplantation of cryopreserved ovarian tissue in a patient with ovarian failure after chemotherapy. N Engl J Med 353, 318–321.

Meirow, D., Ra'anani, H., Shapira, M., Brenghausen, M., Chaim, S.D., Aviel-Ronen, S., Amariglio, N., Schiff, E., Orvieto, R., Dor, J., 2016. Transplantations of frozen-thawed ovarian tissue demonstrate high reproductive performance and the need to revise restrictive criteria. Fertil Steril 106, 467–474.

Oktem, O., Alper, E., Balaban, B., Palaoglu, E., Peker, K., Karakaya, C., Urman, B., 2011. Vitrified human ovaries have fewer primordial follicles and produce less antimüllerian hormone than slow-frozen ovaries. Fertil Steril 95 (8). https://doi.org/10.1016/j.fertnstert.2010.12.057 2661-4.e1Epub 2011 Feb 5. PMID: 21300348.

Parmegiani, L., Accorsi, A., Cognigni, G.E., Bernardi, S., Troilo, E., Filicori, M., 2010. Sterilization of liquid nitrogen with ultraviolet irridation for safe vitrification of human oocytes or embryos. Fertility and Sterility_ 94 (4) September.

Parmegiani, L., Tatone, C., Cognigni, GE., Bernardi, S., Troilo, E., Arnone, A., Maccarini, AM., Di Emidio, G., Vitti, M., Filicori, M., 2014. Rapid warming increases survival of slow-frozen sibling oocytes: a step towards a single warming procedure irrespective of the freezing protocol? Reprod Biomed Online 28 (5), 614–623.

Parmegiani, L., Beilby, K.H., Arnone, A., Bernardi, S., Maccarini, A.M., Nardi, E., Cognigni, G.E.., Filicori, M., 2018. Testing the efficacy and efficiency of a single "universal warming protocol" for vitrified human embryos: prospective randomized controlled trial and retrospective longitudinal cohort study. Journal of assisted reproduction and genetics 35 (10), 1887–1895. https://doi.org/10.1007/s10815-018-1276-4.

Rahimi, G., Isachenko, V., Todorov, P., Tawadros, S., Mallmann, P., Nawaroth, F., Isachenko, E., 2009. Apoptosis in human ovarian tissue after conventional freezing or vitrification and xenotransplantation. Cryo Letters 30 (4), 300–309 PMID: 19789827.

Rimon, E., Cohen, T., Dantes, A., Hirsh, L., Amit, A., Lessing, JB., Freimann, S., Amsterdam, A., Azem, F., 2005. Apoptosis in cryopreserved human ovarian tissue obtained from cancer patients: a tool for evaluating cryopreservation utility. Int J Oncol 27 (2), 345–353. https://doi.org/10.3892/ijo.27.2.345 PMID: 16010414.

Rodriguez-Wallberg, K.A., Tanbo, T., Tinkanen, H., Thurin-Kjellberg, A., Nedstrand, E., Kitlinski, M.L., Macklon, K.T., Ernst, E., Fedder, J., Tiitinen, A, Morin-Papunen, L., Einarsson, S., Jokimaa, V., Hippeläinen, M., Lood, M., Gudmundsson, J., Olofsson, J.I., Andersen, C.Y, 2017. Ovarian tissue cryopreservation and transplantation among alternatives for fertility preservation in the Nordic countries - compilation of 20 years of multicenter experience. Acta obstetricia et gynecologica Scandinavica 95 (9), 1015–1026. https://doi.org/10.1111/aogs.12934.

Rodriguez-Wallberg, K.A., Tanbo, T., Tinkanen, H., Thurin-Kjellberg, A., Nedstrand, E., Kitlinski, M.L., Macklon, K.T., Ernst, E., Jensen, A.K., Macklon, K.T., Fedder, J., Ernst, E., Humaidan, P., Andersen, C.Y, 2017. 86 Successful births and 9 ongoing pregnancies worldwide in women transplanted with frozen-thawed ovarian tissue: focus on birth and perinatal outcome in 40 of these children. J Assist Reprod Genet 34, 325–336.

Rozen, G, Sii, S, Agresta, F, Gook, D, Polyakov, A, Stern, C., 2021. Ovarian tissue grafting: Lessons learnt from our experience with 55 grafts. Reprod Med Biol 20 (3), 277–288. https://doi.org/10.1002/rmb2.12380 PMID: 34262395; PMCID: PMC8254169.

Schallmoser, A., Einenkel, R., Färber, C., Hüren, V., Emrich, N., John, J., Sänger, N., 2023. Comparison of angiogenic potential in vitrified vs. slow frozen human ovarian tissue. Sci Rep 13 (1), 12885. https://doi.org/10.1038/s41598-023-39920-x PMID: 37558708; PMCID: PMC10412559.

Schallmoser, A., Einenkel, R., Färber, C., Emrich, N., John, J., Sänger, N., 2022. The effect of high-throughput vitrification of human ovarian cortex tissue on follicular viability: a promising alternative to conventional slow freezing? Arch Gynecol Obstet. https://doi.org/10.1007/s00404-022-06797-6 Online ahead of print. PMID: 36175682.

Schallmoser, A., Einenkel, R., Färber, C., Hüren, V., Pougin, A., Emrich, N., John, J., Sänger, N., 2023. Cryostorage of human ovarian tissue: evaluating the storage and disposal pattern over a 22-year period in 2475 patients. Reprod Biomed Online 47 (3), 103239. https://doi.org/10.1016/j.rbmo.2023.05.011 Epub 2023 May 27. PMID: 37400319.

Schallmoser, A., Einenkel, R., Färber, C., Sänger, N., 2022. In vitro growth (IVG) of human ovarian follicles in frozen thawed ovarian cortex tissue culture supplemented with follicular fluid under hypoxic conditions. Arch Gynecol Obstet 306 (4), 1299–1311.

Shapira, M, Dolmans, MM, Silber, S, Meirow, D., 2020. Evaluation of ovarian tissue transplantation: results from three clinical centers. Fertil Steril 114 (2), 388–397. https://doi.org/10.1016/j.fertnstert.2020.03.037 Epub 2020 Jun 27. PMID: 32605799.

Shi, Q., Xie, Y., Wang, Y., Li, S., 2017. Vitrification versus slow freezing for human ovarian tissue cryopreservation: a systematic review and meta-anlaysis. Sci Rep 7, 8538. https://doi.org/10.1038/s41598-017-09005-7 Published online 2017 PMCID: PMC5561141 PMID: 28819292.

Silber, S.J., DeRosa, M., Goldsmith, S., Fan, Y., Castleman, L., Melnick, J., 2018. Cryopreservation and transplantation of ovarian tissue: results from ine centerin the USA. J Assist Reprodu Genet 35 (12).

Suzuki, N., Yoshioka, N., Takae, S., Sugishita, Y., Tamura, M., Hashimoto, S., Morimoto, Y., Kawamura, K., 2015. Successful fertility preservation following ovarian tissue vitrification in patients with primary ovarian insufficiency. Hum Reprod 30 (3), 608–615.

Van der Ven, H., Liebenthron, J., Beckmann, M., Toth, B., Korell, M., Krüssel, J., Frambach, T., Kupka, M., Hohl, M.K., Winkler-Crepaz, K., Seitz, S., Dogan, A., Griesinger, G., Häberlin, F., Henes, M., Schwab, R., Sütterlin, M., von Wolff, M., 2016. Ninety-five orthotopic transplantations in 74 women of ovarian tissue after cytotoxic treatment in a fertility preservation network: tissue activity, pregnancy and delivery rates. Hum Reprod 31, 2031–2041.

Wallace, W.H., Kelsey, T.W., Anderson, R.A., 2016. Fertility preservation in pre-pubertal girls with cancer: the role of ovarian tissue cryopreservation. Fertil Steril.

Xiao, Z., Wang, Y., Li, L., Luo, S., Li, S.W., 2010. Needle immersed vitrification can lower the concentration of cryoprotectant in human

ovarian tissue cryopreservation. Fertil. Steril. 94, 2323–2328.
Xiao, Z., Wang, Y., Li, L.L., Li, S.W., 2013. In vitro culture thawed human ovarian tissue: NIV versus slow freezing method. Cryo Letters 34 (5), 520–526 PMID: 24448772.
Zhou, XH., Zhang, D., Shi, J., Wu, YJ., 2016. Comparison of vitrification and conventional slow freezing for cryopreservation of ovarian tissue with respect to the number of intact primordial follicles: A meta-analysis. Medicine (Baltimore) 95 (39), e4095. https://doi.org/10.1097/MD.0000000000004095 PMID: 27684791; PMCID: PMC5265884.

Received 21 December 2023; received in revised form 22 February 2024; accepted 4 March 2024.

RBMO

ELSEVIER

ARTICLE

Evaluation of an artificial intelligence-facilitated sperm detection tool in azoospermic samples for use in ICSI

BIOGRAPHY

Dale Goss is a PhD student at the University of Technology Sydney and a graduate of Stellenbosch University and Monash University. He is a clinical embryologist at IVFAustralia and as a scientific advisor for NeoGenix Biosciences. His research focuses on human embryology, male infertility, and technology in assisted reproduction.

Dale M. Goss[1,2,3,†], Steven A. Vasilescu[1,2,†], Phillip A. Vasilescu[2], Simon Cooke[3], Shannon HK. Kim[3,4], Gavin P. Sacks[1,3,4], David K. Gardner[2,5], Majid E. Warkiani[1,2,6,*]

KEY MESSAGE

This proof-of-concept study shows that an artificial intelligence image analysis tool can drastically improve sperm search times on testicular tissue samples, thus reducing physical strain and fatigue on embryologists and possibly improving the chance of finding spermatozoa. This is a highly translatable clinical tool for the treatment of severe male factor infertility.

ABSTRACT

Research question: Can artificial intelligence (AI) improve the efficiency and efficacy of sperm searches in azoospermic samples?

Design: This two-phase proof-of-concept study began with a training phase using eight azoospermic patients (>10,000 sperm images) to provide a variety of surgically collected samples for sperm morphology and debris variation to train a convolutional neural network to identify spermatozoa. Second, side-by-side testing was undertaken on two cohorts of non-obstructive azoospermia patient samples: an embryologist versus the AI identifying all the spermatozoa in the still images (cohort 1, $n = 4$), and a side-by-side test with a simulated clinical deployment of the AI model with an intracytoplasmic sperm injection microscope and the embryologist performing a search with and without the aid of the AI (cohort 2, $n = 4$).

Results: In cohort 1, the AI model showed an improvement in the time taken to identify all the spermatozoa per field of view ($0.02 \pm 0.30 \times 10^{-5}$s versus 36.10 ± 1.18s, $P < 0.0001$) and improved recall ($91.95 \pm 0.81\%$ versus $86.52 \pm 1.34\%$, $P < 0.001$) compared with an embryologist. From a total of 2660 spermatozoa to find in all the samples combined, 1937 were found by an embryologist and 1997 were found by the AI in less than 1000th of the time. In cohort 2, the AI-aided embryologist took significantly less time per droplet (98.90 ± 3.19 s versus 168.7 ± 7.84 s, $P < 0.0001$) and found 1396 spermatozoa, while 1274 were found without AI, although no significant difference was observed.

Conclusions: AI-powered image analysis has the potential for seamless integration into laboratory workflows, to reduce the time to identify and isolate spermatozoa from surgical sperm samples from hours to minutes, thus increasing success rates from these treatments.

[1] School of Biomedical Engineering, University of Technology Sydney, Sydney, New South Wales, Australia.
[2] NeoGenix Biosciences Pty Ltd, Sydney, New South Wales, Australia.
[3] IVFAustralia, Sydney, New South Wales, Australia.
[4] University of New South Wales, Sydney, New South Wales, Australia.
[5] Melbourne IVF, Melbourne, Victoria, Australia.
[6] Institute for Biomedical Materials & Devices (IBMD), University of Technology Sydney, Sydney, New South Wales, Australia.
† The authors should be considered as joint first authors.

*Corresponding author. E-mail address: majid.warkiani@uts.edu.au (M.E. Warkiani). https://doi.org/10.1016/j.rbmo.2024.103910 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS

Azoospermia
Male infertility
Microdissection testicular sperm extraction
Spermatozoa
Surgical sperm collection

INTRODUCTION

Male infertility is increasing worldwide at an alarming rate, sperm counts having declined by 50% over the past 50 years (*Levine et al., 2023*). Around 30% of cases of human infertility are caused solely by male infertility and 50% of cases are attributed to having male infertility as a contributing factor (*Agarwal et al., 2015*). While assisted reproductive technology has proved to be effective in treating infertile couples, some forms of male infertility remain difficult to treat. Azoospermia, defined as the absence of spermatozoa in centrifuged semen on at least two occasions, is the most severe form of male infertility, affecting 10–20% of infertile men and 1% of the general male population (*Verheyen et al., 2017*; *Wosnitzer et al., 2014*).

Azoospermia can be classified as obstructive and/or non-obstructive. Obstructive azoospermia occurs due to obstruction of the reproductive tract and constitutes 40% of azoospermic cases, while non-obstructive azoospermia (NOA) results from primary, secondary or incomplete/ambiguous testicular failure, which compromises sperm production and constitutes 60% of cases of azoospermia (*Jarow et al., 1989*; *Wosnitzer, et al., 2014*). Patients with obstructive azoospermia can attempt reconstruction (vasovasostomy, vasoepididymostomy or transurethral resection of the ejaculatory duct) when possible, or surgical sperm collection can be performed from the testis via testicular sperm aspiration (TESA), testicular sperm extraction (TESE) or microdissection TESE (mTESE), or from the epididymis via microsurgical epididymal sperm aspiration or percutaneous epididymal sperm aspiration (*Flannigan et al., 2017*; *Schrepferman et al., 2001*). Patients with NOA require sperm extraction from the testis (TESA, TESE or mTESE), and the surgically collected spermatozoa are then used for intracytoplasmic sperm injection (ICSI).

The gold-standard for treating patients with NOA is mTESE, which has a high sperm retrieval rate of up to 64% in suitable patients (*Deruyver et al., 2014*; *Ramasamy et al., 2005*; *Schiff et al., 2005*). Although these rates seem promising, the current manual examination process to find spermatozoa within tissue recovered from mTESE operations is time-consuming and inefficient, typically taking anywhere between 1 and 6 h of laboratory time, and in some cases even up to 14 h (*Mangum et al., 2020*; *Ramasamy et al., 2011*). This extended time is due to the requirement for manual searching through prepared suspensions of testicular tissue with a microscope, before using isolated sperm for ICSI.

The outcome of such searching is heavily dependent upon the complexity and contamination of the suspension provided to the embryologists by the surgeon. Viable sperm are easily overlooked due to variables such as collateral cell density, resulting in a process that is prone to human error, combined with inexperience and fatigue of laboratory staff (*Ramasamy, et al., 2011*). For patients with NOA, overlooking spermatozoa because of human error could wrongly indicate absolute infertility (*Samuel et al., 2016*). Similarly, for extended sperm searches in semen as a diagnostic test or as a last check of the ejaculate before surgery, failure to identify any spermatozoa present could unnecessarily direct patients into surgery. Furthermore, prolonged sample examination procedures can have adverse effects on the viability of the spermatozoa, consequently affecting their potential for fertilization and thus undermining the efforts of sperm searches and the considerable cost and physical strain caused to patients during these procedures (*Ouitrakul et al., 2018*). For individuals with NOA, a more efficient and higher throughput method capable of locating and isolating spermatozoa from the suspension would therefore greatly benefit the clinical workflow of assisting severe forms of male infertility.

Panning through surgically collected sperm samples under a microscope is a form of manual image analysis which machine learning and artificial intelligence (AI) have the potential to automate and improve. Therefore, with preliminary works showing promising results (*Goss et al., 2023*), this study aims to comprehensively assess the use of an assistive convolutional neural network (CNN) AI that was developed and trained to identify spermatozoa in complex tissue suspensions in real time (**FIGURE 1**). Using a YOLOv8 model (Ultralytics, USA), an open-source, high-speed, high-accuracy object detection and image segmentation model, this software works in tandem with an embryologist to instantly identify and alert embryologists to spermatozoa of interest for their assessment from the camera feed mounted into their microscope. The objective of this study was, for cohort 1, to compare the AI with embryologists working without the aid of the AI aid in terms of time, recall and number of spermatozoa found using still images, and then for cohort 2, to run a simulated sperm search with the AI integrated into an ICSI microscope kit, to demonstrate its potential for clinical implementation.

MATERIALS AND METHODS

Ethical approval

Ethical approval for healthy sperm samples was received from the University of Technology Sydney ethics review board (ETH19-3677, approved 12 December 2019), and for the use of discarded testicular tissue samples was received from the IVFAustralia Human Research Ethics Committee (DG01192, approved 2 August 2022) and University of Technology Sydney ethics review board (ETH22-7189, approved 30 September 2022).

Preparation of specifically prepared samples

To generate images for the training dataset prior to access to clinical testicular tissue samples, specifically prepared samples were used for the initial training of the AI model (**FIGURE 1**A). These samples consisted of donor spermatozoa, fingerprick blood and cells from epithelial cell culture lines.

Human semen samples were obtained through ejaculation after 2–5 days of sexual abstinence (*WHO, 2021*). Raw semen samples were left at room temperature for 20 min to allow for liquefaction. Samples were centrifuged for 8 min at 500*g* to separate the sperm pellet from the seminal plasma. Red blood cells (RBC) were obtained from whole-blood specimens within 3 days of collection. The collected blood samples were also resuspended in G-MOPS Plus (Vitrolife, Sweden) medium. Mixed cell suspensions were created to simulate testicular tissues samples containing spermatozoa, RBC, white blood cells (WBC), epithelial cells and C2C12 and THP-1 cells (Sigma-Aldrich, USA). All the cells were mixed in warmed G-MOPS Plus (37°C). Raw semen samples were diluted down to between 1×10^7 and 1×10^8 spermatozoa/ml, the RBC concentration was in the range

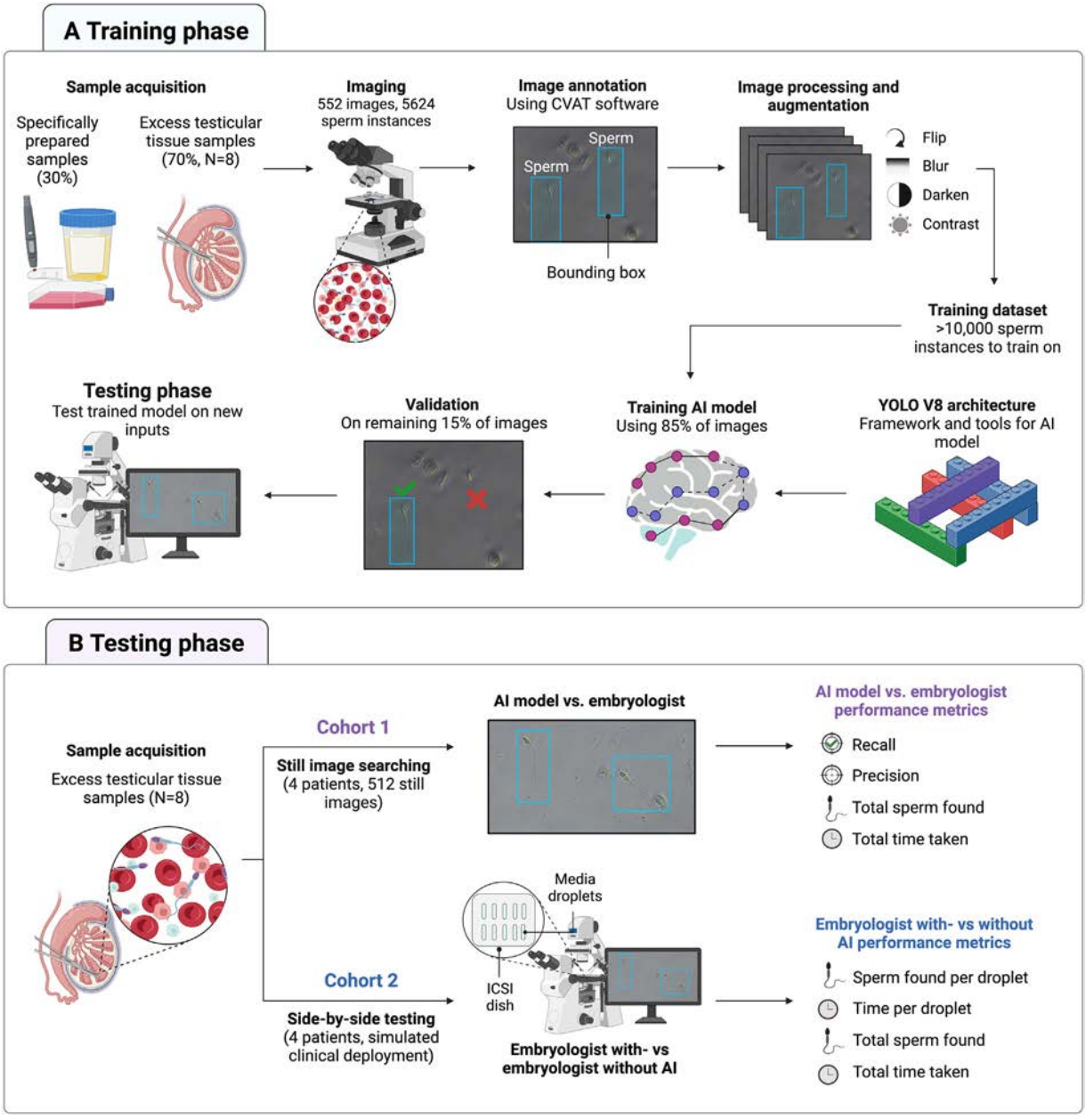


FIGURE 1 Overview of the study phases. (A) The training phase begins with sample acquisition from 30% specifically prepared samples and 70% testicular tissue samples, which are plated in dishes and imaged at a magnification of 200 ×. The images are then annotated using the Computer Vision Annotation Tool (CVAT), creating bounding boxes around all the spermatozoa in each image. Each image is processed and augmented to create a training dataset, which is then used to train a model created using the YOLOv8 architecture and tools. Then 85% of these images are used to train the model and 15% are used to validate the model performance before testing. (B) The testing phase begins with sample acquisition of excess testicular tissue from patients with non-obstructive azoospermia (NOA) and testing of the model's performance on still images (cohort 1, n = 4) versus an embryologist, and side-by-side testing of an embryologist with and without the aid of the artificial intelligence (AI) in a simulated real-world sperm search using an intracytoplasmic sperm injection (ICSI) microscope (cohort 2, n = 4).

2–15 × 10^6 cells/ml (approximated ranges for an mTESE sample), WBC (10 × 10^6 cells/ml; purchased from IQ Biosciences USA) were diluted to a concentration between 5 × 10^5 and 1 × 10^6 cells/ml, and epithelial cells were diluted to a concentration of between 7 × 10^5 and 1 × 10^6 cells/ml.

To add extra complexity, background cells from sperm donors were isolated from donors with high concentrations of background cell populations and cryopreserved until needed. These cells helped to simulate the conditions of poor-quality samples with high levels of collateral cell contamination from surgery and for infertile semen samples with high levels of contamination in the ejaculate.

Testicular biopsy retrieval and processing

Surgical sperm collection was performed in accordance with the routine workflow for each method of sperm collection

(mTESE and TESA) from azoospermic patients scheduled for surgical sperm collection for obstructive azoospermia or NOA. Surgical sperm collections were performed under general anaesthesia, and the samples were immediately placed in a sterile conical tube containing 1 ml of G-MOPS Plus (37°C) and transported to the IVF laboratory.

During mTESE, embryologists search through seminiferous tubules handed to them by the surgeon, with simultaneous further searching by the surgeon for dilated seminiferous tubules. Further samples are then sent to the IVF laboratory for a further search before being placed in 1–2 ml of G-MOPS Plus in a sterile Petri dish under a stereo-microscope to wash off excess blood from the tissue, and then moved to a new Petri dish with 300 μl of G-MOPS Plus. The tissue was gently teased apart using sterile syringes to release potential spermatozoa from the tubules into the surrounding G-MOPS Plus medium. The macerated tissue and large pieces were then removed and placed into a separate tube, and the remaining suspension was used for the sperm search and treatment. In cases whereby imaging and/or testing was not possible on the same day or the following day, the samples were fixed with 4% formalin to preserve their morphological integrity and prevent any microbial growth until use in the study.

To prepare samples for comparison between an AI-enabled sperm search and a sperm search by an embryologist in cohort 1, samples that were recorded having no spermatozoa found in clinical searches were spiked with low concentrations of spermatozoa from semen donors (prepared as described in 'Preparation of Specifically Prepared Samples'). To help create a master count of total spermatozoa in the plated samples, spiked spermatozoa were stained with propidium iodide and washed to remove excess stain before spiking. This was done to help identify the total number of spermatozoa to be found in each sample for comparison with the AI and embryologist performance groups. Samples that had spermatozoa present in the clinics were not spiked with donor semen and were preserved in their clinical state for processing.

Image acquisition and processing

To train the model, specifically prepared samples containing mixtures of spermatozoa, RBC, WBC and epithelial cells from cell culture media were prepared and plated in a similar manner to a clinical sperm search, using 10 long drops of G-MOPS Plus of 2–3 mm in length under OVOIL (Vitrolife, Sweden) in an ICSI dish (Vitrolife, Sweden), and imaged at 200 × magnification using cellSens Imaging Software (Olympus Life Science, Japan; FIGURE 1A). This approach was chosen to initiate training, and once the model's ability to identify spermatozoa was confirmed, clinically obtained testicular tissue samples from eight azoospermic patients (six with NOA and two with obstructive azoospermia) were then used to train the model with more representative backgrounds.

The training dataset comprised 540 images (152 from specifically prepared samples and 388 from testicular tissue samples), containing 5624 unique sperm instances, duplicated and augmented generating at least one augmented copy per image, which resulted in over 10,000 spermatozoa to train the identification function (FIGURE 1A). Synthetic data (duplication) during the model training was used to create more unique images for the model to learn from and is commonly performed to improve dataset fidelity (*Chavez-Badiola et al., 2020*; *Cubuk et al., 2018*; *Trembley et al., 2018*). By creating these flipped and augmented duplicate images, these images can be used in the training process as they may be considered functionally unique to their original copy (Supplementary Figure 1).

Images were annotated using the Computer Vision Annotation Tool (CVAT; Intel, USA) which is open-source software with a web-based interface designed for image and video annotation for computer vision tasks. This software was chosen to create the annotated dataset of images whereby the spermatozoa in these images were annotated with simple bounding boxes (Supplementary Figure 2) that enclose the entire visible spermatozoa including the head and tail. If the spermatozoon is partially occluded it is still bound by a single bounding box encompassing all the visible areas. CVAT was chosen for this purpose due to collaborative annotation from multiple users (including the AI model) as well as the user-friendly interface.

Dataset preparation

The training images were 2456 × 1842px JPG images with 95% compression. Images were saved in JPG format to better reflect real-world environments where images may be sent over a network and require rapid real-time feedback. These were resized to 1664 × 1664px with a black fill. A total of 85% of the images were used for training and 15% reserved for validation of the model's performance after training. Augmentations were applied to all images including duplicates from both the specifically prepared samples and the excess testicular tissue samples to inflate the dataset and make the trained model more robust to variations in microscope camera images, such as compression artefacts, changing focal length or lighting and colour variations.

A vertical flip was applied to each duplicate image, ensuring it was uniquely different from its source, and then with various probabilities a series of augmentation techniques were employed using the Python-based Albumentations library (*Buslaev et al., 2020*). Initially, a blurring effect with a kernel size of 2 × 2 pixels was applied to each image to simulate the effect of slight defocusing. Thereafter, JPG compression was implemented, adjusting the compression quality to a range between 60% and 80%, to mimic the common lossy compression artefacts (features identifiable with the human eye) found in digital imaging. An example of these augmentations is shown in Supplementary Figure 1.

Training of the AI model

Once the dataset of images had been compiled for training, an open-source machine-learning model architecture, YOLOv8 (Ultralytics, USA), was chosen, which provided the framework and tools to train the authors' own model. YOLOv8 was selected as it is a highly performant architecture for real-time object detection tasks, which suits the application of identifying spermatozoa in highly complex tissue samples. YOLOv8 was used with the 'small' size architecture configuration with 225 layers and 1,1166,560 parameters to prioritize minimal inference time (i.e. speed of identifying potential spermatozoa during searching) over potentially greater recall from more parameters (*Jocher et al., 2023*). Further image augmentations were applied by YOLOv8 during the training process, including horizontal flipping, scaling, translation and augmentations to hue, saturation and value. The training setup was restricted to a modest video random access memory of less than 8 GB, which limits the size of the model and training image resolution. Thus,

to maintain a high image resolution required to differentiate fine detail and the desired model size on this set-up, the study used a small batch size of four images being trained in parallel. The model was trained for 300 training iterations or epochs with a learning rate of 0.01. The stochastic gradient descent optimizer was used with 0.937 momentum and 0.005 weight decay.

The trained model was then used to make inferences on unseen, unlabelled images from the 15% of images allocated for validation (FIGURE 1A). The performance of the model was validated on images with a ground truth sperm number showing 85% precision and 78% accuracy after 300 epochs; the model was then considered ready for side-by-side testing against an embryologist as the training dataset is purposely compiled to validate performance on edge-cases and relatively difficult to identify spermatozoa. This approach has been proven to produce robust and unbiased image detection models (*Vabalas et al., 2019*).

Comparison of the AI model versus embryologist performance

Side-by-side testing was split into two cohorts both using immotile spermatozoa for their ability to standardize sperm spatial detection. The first cohort used fixed samples at University of Technology Sydney research laboratories and consisted of comparing the time, recall and precision of sperm detection on still images between the AI model and an embryologist (FIGURE 1B). The AI model was loaded onto a desktop computer (Intel Core i5-10600K CPU @ 4.10 GHz [6 cores] (Intel, USA), RTX 3070 graphics card Zotac Gaming, Hong Kong) and annotated spermatozoa in still images of plated discarded testicular tissue samples (n = 4 NOA patients, 512 images acquired with a total of 2660 spermatozoa to be found) in droplets at 200 × magnification. The embryologist used CVAT to annotate the location of spermatozoa independently in the same images while being timed. Sperm annotations from both the AI model and the embryologist (using the annotation software, CVAT) were then compared with a ground truth of verified sperm labels for each image to attain comparable metrics, i.e. precision, recall, time per field of view (FOV) and total spermatozoa found. Consensus for the ground truth annotation for each image was performed by two scientists independent of the embryologist used to test against the AI model.

Precision is a measure of how many sperm detections are correct, i.e. the ratio of the correctly predicted positive observations to the total number of predictions made, and recall (sensitivity or true positive rate) is a measure of how many of the spermatozoa in an FOV the model finds, i.e. the ratio of correctly predicted positive observations to the total of actual spermatozoa in the FOV. Precision and recall are defined by:

$$Precision = \frac{\sum TP}{\sum (TP + FP)}$$

$$Recall = \frac{\sum TP}{\sum (TP + FN)}$$

where TP is true positives, FP is false positives and FN is false negatives. Potential sperm detections (bounding boxes) with significant overlap (>40% intersection of union) with confirmed spermatozoa were counted as positive detections and those without as negatives. Spermatozoa bordering the edge of an image are often cut off and lack enough information to distinguish them as either positive or negative, so any potential spermatozoa within 2px of the edge of the image were omitted.

For the second cohort, to better simulate real-time clinical deployment, a side-by-side test of the AI comparing the performance of an embryologist with and without the AI was performed. Dishes were plated and prepared testicular tissue samples were added to the dishes in a similar manner to a clinical sperm search, with 10 long drops of G-MOPS Plus of 2–3 mm in length under OVOIL (Vitrolife, Sweden) in an ICSI dish (Vitrolife, Sweden) per patient sample. The embryologist recorded and compared the number of spermatozoa found per droplet for each tissue sample (n = 4 NOA patients) that they processed with (Supplementary Videos 1 and 2) and without AI, as well as the time taken to complete their assessment (FIGURE 1B). No ground truth total sperm number was acquired for each drop or dish and therefore a direct comparison of spermatozoa found per unit time in each drop of medium was compared between using the AI and not using the AI. The embryologist was blinded to the dishes and these were reordered to prevent any memory of the sperm location by the embryologist when performing each search. Confidence of sperm identification functionality was added to the AI whereby the confidence range was indicated by a green (>0.75) and orange (between 0.75 and 0.4) scale (Supplementary Video 1) or a red (>0.75) and blue (between 0.75 and 0.4) scale (Supplementary Videos 2 and 3).

Statistical analysis

All statistical analyses were performed using GraphPad Prism 9.0 (GraphPad Software). Normal distribution was assessed using the Shapiro–Wilk test. The statistical significance of the differences between the groups were tested using the Mann–Whitney U-test as the data were not normally distributed. Two-way analysis of variance was performed to assess the effects of the counting method and group. A value of $P < 0.05$ was considered statistically significant, and the means are expressed with the standard error of the mean (SEM) as a measure of the sample mean estimates.

RESULTS

In the first cohort of this study (n = 4 NOA patients), when assessing the performance of sperm identification from the still images, the AI model showed a dramatic improvement in the time taken to identify the spermatozoa in each FOV, improved recall in identifying spermatozoa and provided a high level of precision (TABLE 1). The AI was able to identify all the spermatozoa within each FOV in significantly less time compared with the trained embryologist, with durations of $0.02 \pm 0.3 \times 10^{-5}$ s versus 36.10 ± 1.18 s, respectively ($P < 0.0001$; TABLE 1). This represents an approximate 99.95% reduction in time per FOV. The AI model demonstrated a significant difference in recall compared with the trained embryologist (91.95 ± 0.81% versus 86.52 ± 1.34%, $P < 0.001$; TABLE 1). The model exhibited a precision of 89.58 ± 0.87%, considering the correct identification of spermatozoa and false positives relative to the control count (TABLE 1). In contrast, the embryologist had a precision of 98.18 ± 0.38% ($P < 0.0001$). Out of a total of 2660 spermatozoa, the embryologist identified 1937, while the AI model detected 1997 (TABLE 1).

In the second cohort of this study (n = 4 NOA patients), a simulated deployment of the AI was performed in a research laboratory whereby the AI was used as an assistive tool to guide the embryologists to identify spermatozoa on an ICSI microscope kit (see Supplementary Videos

TABLE 1 COMPARISON OF AI AND EMBRYOLOGIST SPERM SEARCH PERFORMANCE METRICS

Parameter	Embryologist	AI	*P*-value
Cohort 1 (still images)			
Time per FOV (s)	36.10 ± 1.18	$0.02 \pm 0.3 \times 10^{-5}$	<0.0001[a]
Recall (%)	86.52 ± 1.34	91.95 ± 0.81	0.0006[a]
Precision (%)	98.18 ± 0.38	89.58 ± 0.87	<0.0001[a]
No. of sperm found (from 2660)	1937	1997	N/A
Cohort 2 (side-by-side deployment)			
Time taken per drop (s)	168.7 ± 7.84	98.9 ± 3.19	<0.0001[b]
Total time taken (s)	6749.71	3955.89	N/A
Sperm found per drop	31.85 ± 3.09	34.9 ± 3.43	0.3843[b]
Total no. of sperm found	1274	1396	N/A

Data are presented as the mean ± SEM or total. Between-group differences were tested using a Mann–Whitney *U*-test[a], and variance effects between groups were assessed using two-way analysis of variance.[b]

AI, artificial intelligence; FOV, field of view; N/A, not applicable.

1 and 2). As for cohort 1, the AI-assisted embryologist outperformed the individual assessment of an embryologist across all four samples. The embryologist using the AI took significantly less time to find all the spermatozoa per droplet (98.9 ± 3.19s versus 168.7 ± 7.84s, $P < 0.0001$) and found a total of 1396 spermatozoa, while they found 1274 without the use of the AI (TABLE 1). There was no significant difference in the number of spermatozoa found per droplet for the embryologist using AI versus not using AI although a slight trend of consistently more spermatozoa found was observed (34.9 ± 3.43 versus 31.85 ± 3.09 spermatozoa, respectively).

DISCUSSION

AI image analysis can identify spermatozoa faster and with better recall than an embryologist in still images and significantly faster in a simulated sperm search scenario when integrated into an ICSI microscope. This is the first known application of machine learning AI for surgical sperm searches for the clinical treatment of azoospermia and results in a streamlining of a historically laborious process.

Machine learning is an algorithmic method of data analysis whereby a predictive model is trained to recognize patterns and associations from the input data (*Bannach-Brown et al., 2019*). Supervised machine learning models can be trained on labelled images and/or videos to understand how to predict the labels of unseen data. CNN algorithms are a type of deep-learning model that attempts, through iterative training, to transform input data into the desired output labels. There have been a considerable number of studies on the utility of machine learning and AI-based image analysis on the selection of embryos for the prediction of euploidy status, implantation potential and incidence of miscarriage (*Barnes et al., 2023*; *Diakiw et al., 2022*; *Duval et al., 2023*; *Hariharan et al., 2019*; *Tran et al., 2018*; *VerMilyea et al., 2020*). Studies have also proven the application of machine learning in the selection and assessment of spermatozoa for use in ICSI by tracking spermatozoa correlated with better quality blastocysts (*Joshi et al., 2023*; *Mendizabal-Ruiz et al., 2022*). Furthermore, studies have used images of spermatozoa that have been labelled as normal or abnormally shaped by a professional or stained for DNA integrity; given a sufficient volume and variety of these labelled images, machine learning models have been trained to label the morphology of predicted DNA fragmentation of new, unseen, images of spermatozoa (*McCallum et al., 2019*; *Wang et al., 2019*). Whereas CNN, commonly referred to as AI, have largely looked at spermatozoa in a clear environment, the current authors applied a CNN to complex, processed tissues from testicular sperm retrieval procedures and implemented it in a live video feed for the real-time identification of spermatozoa for use in ICSI.

The application of a computer vision-based machine learning model to identify spermatozoa in real time during sperm searches outperforms embryologists' manual searching in simulated searches using still images in terms of the time taken, recall and sperm count. The biggest noticeable difference is the time reduction, where image analysis is almost instant (0.02 s per FOV) but does not consider clinical tasks such as dish set-up, panning and magnification change, and the collection of identified spermatozoa using a micromanipulator needle. Recall and precision were measured as metrics of both the AI and the embryologists' performance against a ground truth number of spermatozoa per image. The significantly lower time taken to identify spermatozoa per FOV, higher recall and increase in the total number of spermatozoa found show the clear superiority of AI image analysis compared with the eyes and focus of trained embryologists (TABLE 1). Although the AI had a lower precision value than the embryologists in the first cohort, it is worth noting that this is a result of the annotation approach taken when training the AI, and precision values are particularly relevant in applications when the cost of false-positive results is high. For the application of this AI model in sperm searching, the cost of false negatives is much higher, whereby a potential spermatozoon suitable for ICSI could be missed, as opposed to an extra 2 s of an embryologist's attention potentially being wasted in the case of a false-positive result. Recall is, however, essential when the cost of false negatives is high, as is in sperm searches of samples from individuals with NOA.

In the second cohort, testicular tissue samples with supplemented spermatozoa (for better quantification of efficacy) were searched by an embryologist in plated ICSI dishes to better simulate a clinical sperm search on an ICSI kit with and without the aid of the AI (see Supplementary Videos 1 and 2). It was determined that the AI reduced the time taken to identify all the spermatozoa in the droplet by around 50% (TABLE 1), with no drop in the number of spermatozoa identified per drop and a higher total number of spermatozoa identified in total (TABLE 1).

Using an exhaustively trained image analysis model to identify spermatozoa based on tens of thousands of sperm images has clinical utility in directing an embryologist's attention to what the AI deems may be of interest and can thus drastically reduce the time taken or number of manual extended sperm searches when integrated with a micromanipulator microscope. The model

trained in this study is designed to cater for multiple clinics that may have different microscopes, light environments, filters and cameras. These environmental and equipment factors may affect the performance of the AI and have thus been catered for. The image augmentations such as blur, colour variations, focus changes, image saturation and colour balance changes and flipping of images used to train the AI model follow a common strategy in computer vision image analysis whereby these augmentations artificially replicate variant circumstances that may appear in images that were not necessarily widely represented in the training data that comes from a relative few, largely homogenous samples (*Chavez-Badiola et al., 2020*; *Cubuk et al., 2018*; *Trembley et al., 2018*). It is common for microscope images to be slightly blurry or have different lighting conditions and this is replicated in the training data through the authors' choice of augmentations, such that the model is resilient to these conditions. This is another area that with further tuning could improve model performance in the future. The model was also trained using both epididymal and testicular spermatozoa to broaden the sample dataset empowering the AI to broaden target sperm prompting. Importantly, the model can also identify spermatozoa with a broad range of motility, from immotile to hyperactivated, and adjusts and adapts to magnification change and panning in real time (Supplementary Video 3).

The role of this model is not to replace an embryologist, but to be a guide towards spermatozoa of interest, leaving the embryologist to make the final determination on the suitability of a spermatozoon for ICSI. AI can negate the biological limits of human error and observation as well as the effects of fatigue, which have long been a limiting factor to extended sperm searches of heterogenous samples obtained via surgical sperm collection. It is important to remember, however, that the AI is limited to detection within the manually directed FOV, and thus if the embryologist has overlooked an area in the sample, the AI will not be able to detect a spermatozoon without having it within view.

This study was performed solely on immotile spermatozoa for the most accurate quantification for spatially identifying and locating spermatozoa, although the AI identifies motile spermatozoa very well (see Supplementary Video 3), and a true clinical deployment will better prove the clinical utility of the model. This proof-of-concept study demonstrates the potential for AI-assisted sperm searches, both in semen for extended sperm searches and in testicular tissue. While the results of this study are promising, continuing to improve the core dataset and image variety will make the model more robust and adoptable for clinics with significantly different microscope arrangements, as well as achieving a higher level of recall.

The limitation of a simulated sperm search using an ICSI workstation with and without the use of the AI, using samples spiked with spermatozoa, is that it does not consider the time spent confirming the locations of the spermatozoa in the FOV during panning (so as not to re-count or miss spermatozoa). This is a disadvantage of the testing method and might be contributing to the lower difference in time taken per method in cohort 2. Therefore, a robust clinical deployment study has been planned, involving consenting in-treatment patients, whereby embryologists will be able to perform sperm searches with the aid of the AI model.

Furthermore, there is potential for the expansion of this AI to include motility and morphological assessments of identified spermatozoa to help in the choice of spermatozoa for insemination when the spermatozoa outnumber the number of oocytes suitable for injection. Another useful addition to the AI would be a sensitive measure of spermatozoan 'twitching' in these cases. 'Twitching' sperm movement in cases of severe NOA confirms the vitality of the spermatozoa without the need for other interventions to prove sperm vitality such as the hyperosmotic swelling test, which also reduces the time taken when selecting the spermatozoa found.

In conclusion, azoospermia affects 10% of infertile men, with NOA, the most severe form, constituting 60% of these cases (*Verheyen et al., 2017*). The current approaches to recover spermatozoa from men who undergo surgery from this condition are antiquated and potentially detrimental to the quality of the spermatozoa found. This study has successfully demonstrated a proof-of-concept application of an AI image analysis model to drastically reduce the sperm search time in testicular tissue samples in simulated clinical sperm searches. When applying the AI to a simulated real-time search workflow, a 50% reduction in time taken to identify the spermatozoa has been demonstrated. This presents the potential to avoid or at least reduce the negative effect of the extended exposure of spermatozoa to biopsied testicular tissue containing a host of molecules capable of reducing sperm viability. By applying this approach with further development and ergonomic optimization, the authors believe it could result in a standardized and more efficient workflow, greatly improving the current processing procedure of all surgically retrieved samples and azoospermic ejaculates by increasing access to treatment for azoospermia and reducing staff time required, as well as increasing sample coverage to ultimately increase chances of finding spermatozoa.

DATA AVAILABILITY

The data that has been used is confidential.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support and facilitation provided by IVFAustralia Eastern Suburbs doctors Dr Jeffrey Persson and Dr Shadi Khashaba, in providing access to patients undergoing treatment for severe male factor infertility, and the technical support and input from the embryology team. M.E.W. would like to acknowledge the support of the Cancer Institute New South Wales through the Career Development Fellowship (2021/CDF1148).

AUTHOR CONTRIBUTIONS

D.M.G., S.A.V., P.A.V., S.C., S.H.K.K., D.K.G. and M.E.W. designed and conceptualized the study. D.M.G., S.A.V. and P.A.V. were responsible for the data acquisition. P.V. and S.A.V. designed and trained the AI model. D.M.G., S.H.K.K. and S.A.V. facilitated clinical sample acquisition. D.M.G., S.A.V. and P.A.V. drafted the manuscript and all the authors critically revised and finally approved it. All authors have agreed to be accountable for the academic integrity and accuracy of the research.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.rbmo.2024.103910.

REFERENCES

Agarwal, A., Mulgund, A., Hamada, A., Chyatte, M.R., 2015. A unique view on male infertility around the globe. Reproductive biology and endocrinology 13, 1–9.

Bannach-Brown, A., Przybyła, P., Thomas, J., Rice, A.S., Ananiadou, S., Liao, J., Macleod, M.R., 2019. Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. Systematic reviews 8, 1–12.

Barnes, J., Brendel, M., Gao, V.R., Rajendran, S., Kim, J., Li, Q., Malmsten, J.E., Sierra, J.T., Zisimopoulos, P., Sigaras, A., 2023. A non-invasive artificial intelligence approach for the prediction of human blastocyst ploidy: A retrospective model development and validation study. The Lancet Digital Health 5, e28–e40.

Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A., 2020. Albumentations: fast and flexible image augmentations. Information 11, 125.

Chavez-Badiola, A., Flores-Saiffe-Farías, A., Mendizabal-Ruiz, G., Drakeley, A.J., Cohen, J., 2020. Embryo Ranking Intelligent Classification Algorithm (ERICA): artificial intelligence clinical assistant predicting embryo ploidy and implantation. Reproductive BioMedicine Online 41, 585–593.

Cubuk E.D., Zoph B., Mane D., Vasudevan V., Le Q.V., 2018. Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv:1805.09501.

Deruyver, Y., Vanderschueren, D., Van der Aa, F., 2014. Outcome of microdissection TESE compared with conventional TESE in non-obstructive azoospermia: a systematic review. Androl 2, 20–24.

Diakiw, S., Hall, J., VerMilyea, M., Amin, J., Aizpurua, J., Giardini, L., Briones, Y., Lim, A., Dakka, M., Nguyen, T., 2022. Development of an artificial intelligence model for predicting the likelihood of human embryo euploidy based on blastocyst images from multiple imaging systems during IVF. Human Reproduction 37, 1746–1759.

Duval, A., Nogueira, D., Dissler, N., Maskani Filali, M., Delestro Matos, F., Chansel-Debordeaux, L., Ferrer-Buitrago, M., Ferrer, E., Antequera, V., Ruiz-Jorro, M., 2023. A hybrid artificial intelligence model leverages multi-centric clinical data to improve fetal heart rate pregnancy prediction across time-lapse systems. Human Reproduction 38, 596–608.

Flannigan, R., Bach, P.V., Schlegel, P.N., 2017. Microdissection testicular sperm extraction. Translational Androl. and Urology. 6, 745.

Goss, D., Vasilescu, S., Vasilescu, P., Sacks, G., Gardner, D., Warkiani, M., 2023. O-136 Artificial intelligence to assist in surgical sperm detection and isolation. Hum. Reproduction. 38, dead093. 163.

Hariharan, R., He, P., Meseguer, M., Toschi, M., Rocha, J.C., Zaninovic, N., Malmsten, J., Zhan, Q., Hickman, C., 2019. Artificial intelligence assessment of time-lapse images can predict with 77% accuracy whether a human embryo capable of achieving a pregnancy will miscarry. Fertility and Steril. 112, e38–e39.

Jarow, J.P., Espeland, M.A., Lipshultz, L.I., 1989. Evaluation of the azoospermic patient. The J. of Urology. 142, 62–65.

Jocher G., Chaurasia A., Qiu J. YOLO by Ultralytics. 2023. Ultralytics, GitHub.

Joshi, K., Simbulan, R.K., Rajah, A.M., Burd, G., Gupta, S., Behr, B., Guarnaccia, M., Singh, G., 2023. A proof-of-concept prospective study of applying artificial intelligence for sperm selection in the IVF laboratory. Reproductive BioMedicine Online, 103329.

Levine, H., Jørgensen, N., Martino-Andrade, A., Mendiola, J., Weksler-Derri, D., Jolles, M., Pinotti, R., Swan, S.H., 2023. Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. Hum. Reproduction Update 29, 157–176.

Mangum, C.L., Patel, D.P., Jafek, A.R., Samuel, R., Jenkins, T.G., Aston, K.I., Gale, B.K., Hotaling, J.M., 2020. Towards a better testicular sperm extraction: novel sperm sorting technologies for non-motile sperm extracted by microdissection TESE. Translational Androl. and Urology. 9, S206.

McCallum, C., Riordon, J., Wang, Y., Kong, T., You, J.B., Sanner, S., Lagunov, A., Hannam, T.G., Jarvi, K., Sinton, D., 2019. Deep learning-based selection of human sperm with high DNA integrity. Communications Biology 2, 250.

Mendizabal-Ruiz, G., Chavez-Badiola, A., Figueroa, I.A., Nuño, V.M., Farias, A.F-S., Valencia-Murilloa, R., Drakeley, A., Garcia-Sandoval, J.P., Cohen, J., 2022. Computer software (SiD) assisted real-time single sperm selection associated with fertilization and blastocyst formation. Reproductive BioMedicine Online 45, 703–711.

Ouitrakul, S., Sukprasert, M., Treetampinich, C., Choktanasiri, W., Vallibhakara, S.A-O., Satirapod, C., 2018. The Effect of Different Timing after Ejaculation on Sperm Motility and Viability in Semen Analysis at Room Temperature. J. of the Méd Association of Thail. 101.

Ramasamy, R., Reifsnyder, J.E., Bryson, C., Zaninovic, N., Liotta, D., Cook, C-A., Hariprashad, J., Weiss, D., Neri, Q., Palermo, G.D., 2011. Role of tissue digestion and extensive sperm search after microdissection testicular sperm extraction. Fertil. and Steril. 96, 299–302.

Ramasamy, R., Yagan, N., Schlegel, P.N., 2005. Structural and functional changes to the testis after conventional versus microdissection testicular sperm extraction. Urology 65, 1190–1194.

Samuel, R., Badamjav, O., Murphy, K.E., Patel, D.P., Son, J., Gale, B.K., Carrell, D.T., Hotaling, J.M., 2016. Microfluidics: The future of microdissection TESE? Systems Biology in Reproductive Medicine 62, 161–170.

Schiff, J.D., Palermo, G.D., Veeck, L.L., Goldstein, M., Rosenwaks, Z., Schlegel, P.N., 2005. Success of testicular sperm injection and intracytoplasmic sperm injection in men with Klinefelter syndrome. The J. of Clinical Endocrinology & Metabolism. 90, 6263–6267.

Schrepferman, C.G., Carson, M.R., Sparks, A.E., Sandlow, J.I., 2001. Need for sperm retrieval and cryopreservation at vasectomy reversal. The J. of Urology 166, 1787–1789.

Tran, A., Cooke, S., Illingworth, P., Gardner, D., 2018. Artificial intelligence as a novel approach for embryo selection. Fertil. and Steril. 110, e430.

Tremblay J., Prakash A., Acuna D., Brophy M., Jampani V., Anil C., To T., Cameracci E., Boochoon S., Birchfield S. Training deep networks with synthetic data: Bridging the reality gap by domain randomization Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2018, pp. 969-977.

Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PloS one 14, e0224365.

Verheyen, G., Popovic-Todorovic, B., Tournaye, H., 2017. Processing and selection of surgically-retrieved sperm for ICSI: a review. Basic and Clinical Androl. 27, 1–10.

VerMilyea, M., Hall, J., Diakiw, S., Johnston, A., Nguyen, T., Perugini, D., Miller, A., Picou, A., Murphy, A., Perugini, M., 2020. Development of an artificial intelligence-based assessment model for prediction of embryo viability using static images captured by optical light microscopy during IVF. Hum. Reproduction. 35, 770–784.

Wang, Y., Riordon, J., Kong, T., Xu, Y., Nguyen, B., Zhong, J., You, J.B., Lagunov, A., Hannam, T.G., Jarvi, K., 2019. Prediction of DNA integrity from morphological parameters using a single-sperm DNA fragmentation index assay. Advanced Science 6, 1900712.

WHO, 2021. Laboratory manual for the examination and processing of human semen, Sixth edn World Health Organization.

Wosnitzer, M., Goldstein, M., Hardy, M.P., 2014. Review of azoospermia. Spermatogenesis. 4, e28218.

Received 7 September 2023; received in revised form 31 January 2024; accepted 9 February 2024.

RBMO

REVIEW

Role of genetic analysis of products of conception and PGT in managing early pregnancy loss

BIOGRAPHY

William H. Kutteh, MD, PhD, HCLD, earned his PhD in Molecular Cell Biology and Immunology and is Certified in Obstetrics and Gynecology and Reproductive Endocrinology by the American Board of Obstetrics and Gynecology. He is a Professor at the University of Tennessee and Director of the Recurrent Pregnancy Loss Center at Fertility Associates of Memphis, USA.

William H. Kutteh[1,2,*], Ralph S. Papas[3], Melissa K. Maisenbacher[1], Elias M. Dahdouh[4]

KEY MESSAGE

Combining the results from the standard recurrent pregnancy loss evaluation with 24-chromosome microarray testing on miscarriage tissue will provide a probable explanation for over 90% of patients. This method of evaluation may help to identify those patients expected to benefit from expectant management versus PGT.

ABSTRACT

This article considers the addition of comprehensive 24-chromosomal microarray (CMA) analysis of products of conception (POC) to a standard evaluation for recurrent pregnancy loss (RPL) to help direct treatment towards expectant management versus IVF with preimplantation genetic testing for aneuploidies (PGT-A). The review included retrospective data from 65,333 miscarriages, a prospective evaluation of 378 couples with RPL who had CMA testing of POC and the standard workup, and data from an additional 1020 couples who were evaluated for RPL but did not undergo CMA testing of POC. Aneuploidy in POC explained the pregnancy loss in 57.7% (218/378) of cases. In contrast, the full RPL evaluation recommended by the American Society for Reproductive Medicine identified a potential cause in only 42.9% (600/1398). Combining the data from the RPL evaluation and the results of genetic testing of POC provides a probable explanation for the loss in over 90% (347/378) of women. Couples with an unexplained loss after the standard evaluation with POC aneuploidy accounted for 41% of cases; PGT-A may be considered after expectant management. Conversely, PGT-A would have a limited role in those with a euploid loss and a possible explanation after the standard workup. Categorizing a pregnancy loss as an explained versus unexplained loss after the standard evaluation combined with the results of CMA testing of POC may help identify patients who would benefit from expectant management versus PGT-A.

INTRODUCTION

Miscarriage has been described as one of the most common complications of pregnancy. The condition when a pregnancy loss recurs — recurrent pregnancy loss (RPL) — affects 2–4% of reproductive-aged women worldwide (*Dimitriadis et al., 2020*). The American College of Obstetrics and Gynecology (*ACOG, 2023*), the American Society for Reproductive Medicine (ASRM) (*Practice Committee of the American Society for Reproductive 2012*) and the European Society of Human Reproduction and Embryology (ESHRE) (*RPL et al., 2023*) define RPL as a disease, distinct from

[1] Natera, Inc., San Carlos, CA, USA.
[2] Department of Obstetrics and Gynecology, University of Tennessee Health Sciences Center, Memphis, TN, USA.
[3] Department of Obstetrics and Gynecology, University of Balamand, Beirut, Lebanon.
[4] ART Center, CHU Sainte-Justine, Department of Obstetrics and Gynecology, Université de Montréal, Montreal, Canada.

*Corresponding author. E-mail address: wkutteh50@gmail.com (W.H. Kutteh). https://doi.org/10.1016/j.rbmo.2023.103738

Declaration: W.K.H. undertakes clinical research with Natera. M.K.M. is employed by Natera with the option to hold stock. The other authors report no financial or commercial conflicts of interest.

KEY WORDS

Aneuploidy IVF
Comprehensive chromosomal microarray
Preimplantation genetic testing
Products of conception testing
Recurrent pregnancy loss

infertility, defined by two or more failed clinical pregnancies before 20 weeks' gestation. Maternal age and number of previous miscarriages are important factors in predicting future losses (*Lund et al., 2012*).

ASRM recommended evaluations for RPL include: a consideration of karyotypes in both partners to look for chromosomal rearrangements; lupus anticoagulant, anticardiolipin antibodies, and anti-β_2-glycoprotein-1 antibodies to look for antiphospholipid syndrome (*Kutteh, 2014*); imaging of the uterine cavity by 3D saline-infused ultrasonography or hysterosalpingography, or hysteroscopy, to look for congenital and acquired uterine anomalies; and blood concentrations of prolactin, thyroid stimulating hormone and haemoglobin A_{1c} to look for hormonal imbalances (*Practice Committee of the American Society for Reproductive, 2012*). Increasingly, clinicians include evaluations for uterine infections and sperm DNA fragmentation (*ACOG, 2023*; *Practice Committee of the American Society for Reproductive 2012*; *RPL et al., 2023*). When all these evaluations have been completed, fewer than 45% of all patients have a possible explanation for their RPL (*Jaslow et al., 2010*, *Popescu et al., 2018*). Here, the current authors are extending their previous report (*Maisenbacher et al., 2019*) to include the results from more than 65,000 singleton products of conception (POC), over 55% of the pregnancy losses being explained by chromosomal abnormalities in the fetus.

Based on these observations, a new algorithm has been proposed for the evaluation of RPL, with modifications to the ASRM evaluation. The proposal modifies the existing guidelines in two ways: (i) it adds to the standard evaluation a comprehensive 24-chromosomal analysis (CCA) of the miscarriage tissue after the second documented pregnancy loss; and (ii) it eliminates the karyotype analysis on the parents (*Brezina PR, 2013*; *Papas and Kutteh, 2020*; *Popescu et al., 2018*). Using this strategy, a probable or definite cause will be identified in over 90% of all miscarriages in couples with RPL. When the miscarriage is aneuploid and the reason for the loss is explained, only 25% of cases will have an additional abnormal finding on the standard ASRM/ESHRE workup. When the result is an unbalanced chromosome rearrangement, parental karyotype analysis should be performed, along with genetic counselling. If the CCA indicates a euploid loss, and the explanation for the loss is unknown, the ASRM evaluation is recommended and will reveal an abnormal finding in about 80% of couples (*Dahdouh and Kutteh, 2021*; *Popescu et al., 2018*). This new strategy is projected to result in cost savings for the healthcare system, mainly in the private practices in USA (*Popescu et al., 2018*).

Preimplantation genetic testing (PGT) is a form of genetic testing that requires IVF and embryo biopsy and therefore involves qualified and experienced laboratory staff and an additional expense. It can be considered potentially therapeutic in specific situations. PGT can be used to identify several different types of structural error, primarily structural chromosomal rearrangements (PGT-SR), aneuploidies (PGT-A) and monogenic/single-gene conditions (PGT-M) (*Dahdouh et al., 2015a*). PGT-SR involves a targeted approach that is typically used when known chromosomal rearrangements (e.g. balanced translocation) are present in the parental genomes. The role of PGT-SR in RPL has been well defined in situations with the presence of a balanced parental chromosome rearrangement diagnosed by karyotyping and in the presence of an unbalanced translocation or inversion detected on genetic testing of POC (*Committee et al., 2020*; *Viotti, 2020*). The most recent publications reveal improved clinical and live birth rates (LBR) per pregnancy, as well as decreased miscarriage rates for translocation carriers, when CCA is applied on trophectoderm cells from blastocyst-stage biopsies (*Huang et al., 2019*).

PGT-A, on the other hand, is a technology that was developed to screen for chromosomal aneuploidies that arise spontaneously in embryos. Few studies have specifically investigated the role of PGT-A in RPL. ESHRE advises against its use for couples with RPL without a genetic cause (*Committee et al., 2020*; *RPL et al., 2023*). PGT-A has many limitations including a high cost, a risk of having no euploid embryo to transfer and the presence of embryos with mosaic results (*Dahdouh et al., 2022*). A large prospective randomized controlled trial (RCT) of the use of PGT-A (*Munne et al., 2019*) was conducted using next-generation sequencing (NGS); the investigators reported no advantage for the use of PGT-A versus morphology alone for embryo selection for IVF. Furthermore, in a more recent RCT (*Yan et al. 2021*) performed in women younger than 37 years undergoing IVF for infertility, PGT-A showed an equivalent cumulative LBR in a 1-year period compared with standard morphology. Other studies have shown improved clinical outcomes in the PGT-A group compared with a control group, including a higher LBR per embryo transfer and per treatment cycle, a reduced multiple pregnancy rate and reduced pregnancy loss rate (*Sacchi et al.. 2019*; *Sanders et al.. 2021*). These data, along with previous reports, show clearly that the ideal group of patients who might benefit from the use of PGT-A are patients older than 35 years where two or more blastocysts are available for biopsy (*Dahdouh, 2021*; *Dahdouh et al., 2015b*; *Sacchi et al., 2019*).

The best option for patients when considering expectant management versus PGT-A in RPL has been unclear (*Dahdouh et al., 2021*; *Papas and Kutteh, 2020*). Currently, no recommendation for PGT-A in RPL exists from major societies (*Preimplantation Genetic Testing: ACOG Committee Opinion, Number 799 2020*; *RPL et al., 2023*). The retrospective, multicentre SART-CORS (Society for Assisted Reproductive Technology Clinic Outcome Reporting System) database is the largest study to date assessing the role of PGT-A in RPL (*Bhatt et al., 2021*); women using PGT-A had an improvement in LBR after frozen embryo transfer (*Bhatt et al., 2021*).

The new proposed evaluation aimed to clarify which patients with RPL might benefit from PGT-A, if any. The authors sought to evaluate their recent and previous data when using this new evaluation strategy, with the inclusion of the standard evaluation of RPL plus comprehensive 24-chromosomal microarray (CMA) testing on the POC to determine the frequency of recurrent miscarriages that remain *truly unexplained* (*Dahdouh, 2021*; *Dahdouh and Kutteh, 2021*; *Papas and Kutteh, 2020*). Other groups have performed long-term follow-up studies indicating that the prognosis for these patients is very good and can be predicted based on the age of the female partner and the number of prior losses (*Lund et al., 2012*). The authors thus wanted to determine if this classification system, combining an explained versus unexplained workup with euploid versus aneuploid miscarriage, would assist clinicians in determining those patients with RPL who might benefit from expectant management versus PGT (*Dahdouh and Kutteh, 2021*; *Papas and Kutteh, 2020*).

MATERIALS AND METHODS

The prospective study of RPL patients was approved by the Institutional Review Board for exempt status in 2010 as the research involved the collection and study of existing data recorded by the investigators in such a manner that the subjects could not be identified directly or indirectly through identifiers linked to the subjects. The retrospective review of POC CMA was granted a waiver of the requirement for documentation of informed consent by an Institutional Review Board under 45 CFR 46.104(d)(4) (Salus IRB, 19040-05, 13 March 2023).

Prospective study of RPL patients

Participants eligible for inclusion were patients with RPL referred to the Fertility Associates of Memphis from 2014 to 2022. A total of 378 women aged 20–45 years old diagnosed with RPL were eligible, regardless of socioeconomic status or race (*Papas and Kutteh, 2020*). These 378 had undergone both the full ASRM evaluation and PGT-A of their miscarriage tissue. This group of prospective patients is referred to in this review as the prospective group of 378 RPL patients, As a part of an ongoing investigation of the causes of RPL, all patients were encouraged to have a full evaluation as recommended by the ASRM (*Practice Committee of the American Society for Reproductive, 2012*). All testing was performed when the women were not pregnant and at least 6 weeks remote from a miscarriage. The participants included had experienced at least two clinical pregnancy losses, had undergone a full RPL workup as recommended by the ASRM, and had had a CCA with fetal results from POC following their second or subsequent pregnancy loss. Patients were excluded from the analysis if they had multiple gestations, ectopic pregnancies, biochemical pregnancy losses, losses beyond 20 gestational weeks or pregnancy terminations.

An earlier group 1020 RPL patients had undergone complete testing for RPL but were evaluated before CMA testing of POC had become routine clinical practice. The results from this group of 1020 RPL patients have previously been published (*Jaslow et al., 2010*). This group of 1020 women with RPL were evaluated prospectively. In some of the analysis in this manuscript, for example FIGURE 1, where the causes of RPL were analysed based on the ASRM recommendations, the 1020 RPL patients reported on by Jaslow and colleagues have been combined with the 378 patients evaluated by Papas and Kutteh described above, to give a total of 1398 patients with RPL. Testing was based on ASRM (*Practice Committee of the American Society for Reproductive, 2012*) recommendations that identified definite or probable causes of RPL (*Jaslow et al., 2010*). The evaluation for RPL included all the following:

- Parental karyotypes.
- Evaluation of the uterine anatomy by hysterosalpingography, hysteroscopy or sonohysterography (*Jaslow and Kutteh, 2013*).
- Autoimmune factors: lupus anticoagulant, anticardiolipin and anti-β_2-glycoprotein antibody test results confirmed by repeat testing at least 6 weeks later (*Kutteh, 2014*).
- Endocrine factors including thyroid function, haemoglobin A_{1c} and prolactin.

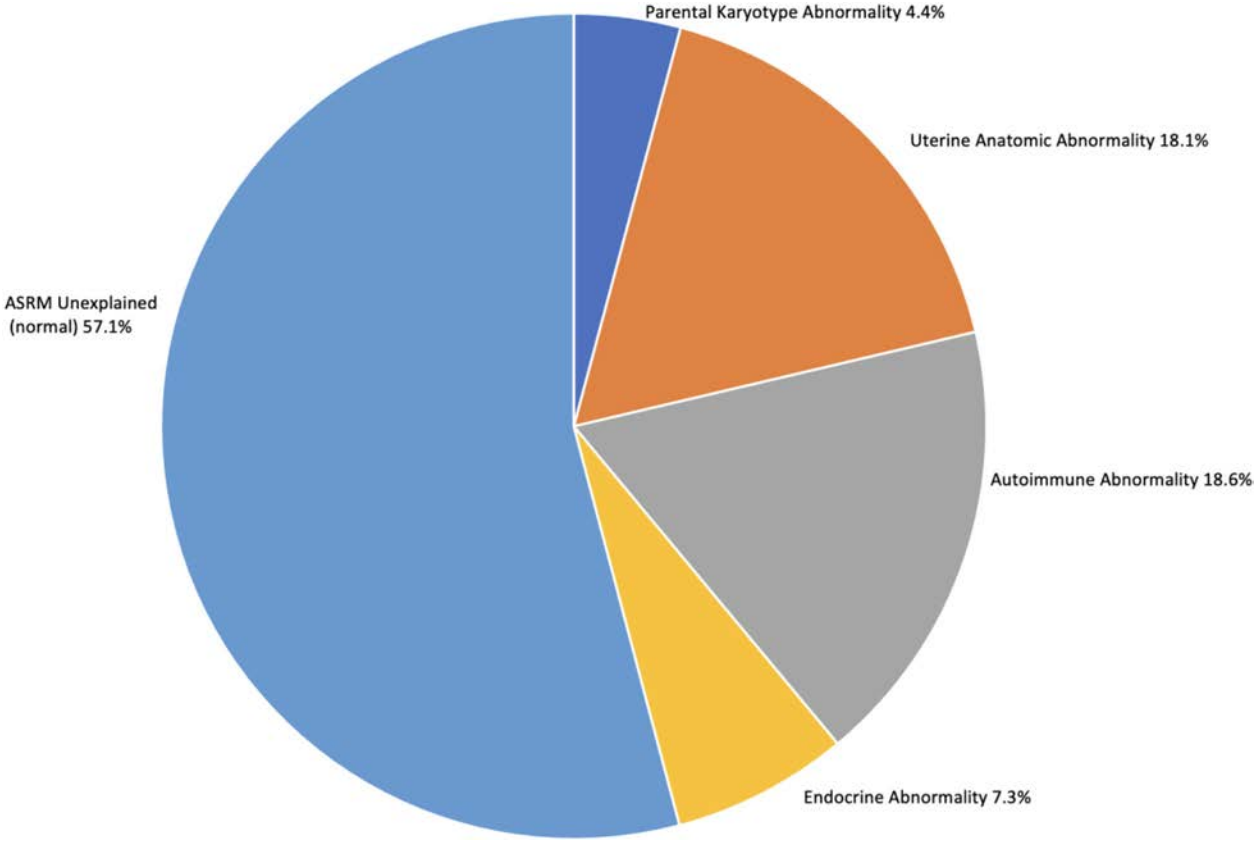


FIGURE 1 Results of the American Society for Reproductive Medicine (ASRM) evaluation of 1398 couples with recurrent pregnancy loss, comprising 1020 and 378 couples from the reports of Jaslow and colleagues (*Jaslow et al., 2010*) and *Papas and Kutteh (2020)*, respectively. The total percentages add up to more than 100% because some patients had more than one abnormality. Normal results (unexplained) were found in 798/1398 (57.1%) of patients.

– CMA of miscarriage tissue, genotyped using HumanCytoSNP-12 BeadChip microarrays, which contain more than 300,000 probes covering all 24 chromosomes with a median distance between probes of 6500 base pairs (Illumina, USA) (*Levy et al., 2014*). The comparison of the single-nucleotide polymorphism (SNP) identities between the maternal and POC data was used to identify DNA copy number, maternal cell contamination (MCC), the parental origin of the aneuploidy, and unbalanced chromosome segments using the previously described proprietary Parental Support TM algorithm (*Johnson et al., 2010*) (Anora, Natera, USA).

Retrospective review of POC CMA

The retrospective review consisted of 65,333 fresh singleton POC specimens received by Natera for evaluation by 24-CMA over 13 years (April 2010 – March 2023). These data expand the authors' previous publication that included 26,101 cases of fresh singleton POC specimens (*Maisenbacher et al., 2019*) and an abstract presented to ASRM that reported on 63,277 cases of fresh singleton POC specimens (*Maisenbacher and Kutteh, 2020*). A maternal blood sample was requested with each specimen to rule out MCC.

Clinical information including maternal age and gestational age at the time of the loss was obtained from information supplied on the test requisition and all samples were de-identified before review. Patients were excluded if they had multiple gestations, ectopic pregnancies, donor oocytes or losses beyond 20 gestational weeks or if no maternal age was available. Samples were classified as 'abnormal' or 'normal' based on the presence or absence of chromosome abnormalities. All terminal copy number variants of 1 Mb of more and interstitial variants over 5 Mb were reported. The abnormalities were categorized as single aneuploidy, complex aneuploidy, deletions/duplications, triploidy and other abnormalities.

RESULTS

Adding POC testing to the ASRM RPL workup explains over 90% of losses

In women who underwent a comprehensive RPL workup as defined by the ASRM, only 600/1398 (42.9%) had an abnormality identified that could be a potential contributing factor for RPL (FIGURE 1). These findings combine data from the authors' previous studies that evaluated abnormalities in a sample population of 1020 women with RPL (*Jaslow et al., 2010*) with more recent data collected from the same clinic on a population of 378 couples with RPL (*Papas and Kutteh, 2020*). The combination of these two datasets (1020 patients plus 378 patients) results in a total population of RPL patients of 1398.

The frequencies of specific abnormalities identified by the ASRM RPL workup (parental karyotype, anatomical, endocrine and autoimmune) are reported in FIGURE 1. Only 42.9% (600/1398) of patients had a possible explanation for the pregnancy loss based on the ASRM workup alone. More than 57% (798/1398) of couples had a normal evaluation and were classified as unexplained RPL.

There were 378 women in this study who undergone CMA of the miscarriage tissue at the time of their second or subsequent miscarriage in addition to the standard ASRM workup for RPL (*Papas and Kutteh, 2020*). The majority of losses (218/378, 57.7%) had an abnormal (aneuploidy plus unbalanced translocation or inversion) POC CMA (FIGURE 2, left-hand pie chart). In women with an aneuploid loss, the majority (163/218, 74.8%) had a normal (unexplained) RPL workup (data not shown). On the other hand, as previously reported, 80% of women who had a normal POC CMA were identified with an abnormality (explained) in the ASRM RPL workup (*Popescu et al., 2018*). Overall, 347/378 (91.8%) had an abnormal POC CMA and/or an abnormal (explained) ASRM RPL workup (data not shown) that would provide a possible or definite cause for the loss (FIGURE 2, right-hand pie chart). Thus, only 31 out of 378 (8.2%) women who were evaluated in this study had a pregnancy loss without a potential explanation.

When the ASRM evaluation was performed on the 378 women with RPL, the majority of women with a normal ASRM evaluation were found to have POC with aneuploidy (*Papas and Kutteh, 2020*; *Popescu et al., 2018*). The finding of euploid POC was linked with a higher probability of finding an abnormality (explained) on the ASRM workup. Conversely, the finding of aneuploid POC was associated with a lower than average chance of finding an abnormality on the ASRM evaluation.

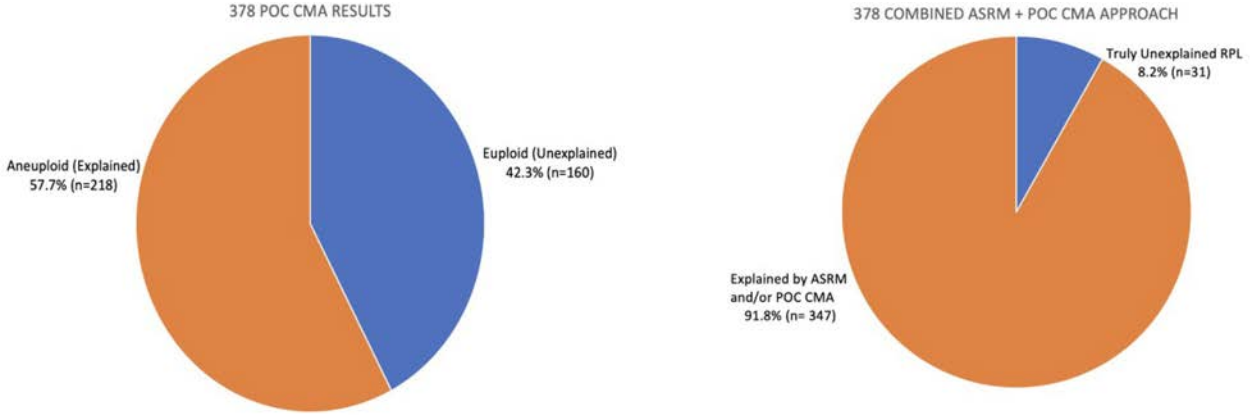


FIGURE 2 Results of the recurrent pregnancy loss (RPL) evaluation when combining 24-chromosomal microarray (CMA) analysis of products of conception (POC) after the second or subsequent loss with a modified American Society for Reproductive Medicine (ASRM) evaluation (without parental chromosome analysis). A total of 91.8% (347/378) couples had a possible or proven explanation for their loss (right-hand pie chart) (*Papas and Kutteh, 2020*). When using only CMA on POC, 57.7% (218/378) of pregnancy losses could be explained by an aneuploidy (left-hand pie chart).

TABLE 1 RESULTS OF 24-CHROMOSOME MICROARRAY TESTING OF PRODUCTS OF CONCEPTION FROM 65,333 PATIENTS WITH A SINGLETON MISCARRIAGE

Parameter	Number	Percentage
POC testing (*n* = 65,333)		
Fetal tissue results	54,912	84.0
Normal fetal results	23,361	42.5
Abnormal fetal results	31,551	57.5
Maternal cell contamination	10,071[a]	15.4
Uninformative	350	0.5
Abnormal results (*n* = 31,551)		
Single aneuploidy	22,912	72.6
Complex aneuploidy	3,559	11.3
Triploidy	3,986	12.6
Deletion/duplication	743	2.4
Other	351	1.1

[a] In a total of 812 cases maternal cell contamination could not be differentiated from normal female.
POC, products of conception.

The majority of miscarriages are caused by aneuploidy

A total of 65,333 POC specimens were tested from singleton intrauterine pregnancies for which maternal age and gestational age were available and the gestational age was less than 20 weeks. Complete data on fetal versus MCC results were available for 64,521 (98.8%) of these. Fetal results were obtained for 54,912 (84%) cases (TABLE 1). The remaining 9,609 (14.9%) cases were excluded from the analysis due to MCC (*n* = 9,259, 14.4%) or incomplete results (*n* = 350, 0.5%). CMA testing identified aneuploidy in 57.5% cases with fetal results (31,551/54,912; TABLE 1). Mean gestational age at the time of the loss was 9.2 weeks (range 4.0–19.9 weeks). The frequency of obtaining a fetal result increased with gestational age, ranging from 53.2% at less than 5 gestational weeks to 97.7% at more than 13 gestational weeks (TABLE 2).

The SART maternal age groups were used to categorize the data from all cases with fetal results (*n* = 54,912). The mean maternal age at the time of the pregnancy loss was 34.0 years (range 18–50 years). The aneuploidy rate varied based on the maternal age at the time of the pregnancy loss from a low of 48.5% (5,833/12,024) in women less than 30 years old to a high of 76.4% (1360/1780) in women over age 42 years (FIGURE 3).

The gestational age at the time of the pregnancy loss appears to be greater in younger women compared with older women based on the maternal age at the time of the loss (FIGURE 4). The mean gestational age at the time of the pregnancy loss in women under age 30 years was 10.4 weeks. In women aged 35–37 years at the time of the pregnancy loss, the mean gestational age at the time of the loss was 8.7 weeks, and in women over age 42 years the pregnancy losses occurred even earlier, at on average 8.4 weeks.

Single aneuploidies were the most common aneuploidies identified (22,912/31,551, 72.6%; TABLE 1). Trisomy was identified in 19,405/31,551 (61.5%) while monosomies were identified in 12.3% (3865/31,551). Mosaic monosomy/trisomy was identified in two cases. Triploidy was identified in 3986/31,551 (12.6%) and varied in frequency based on the maternal age range (FIGURE 5). Triploidy accounted for over 20% of the aneuploidies reported in women under 30 years old, while in women over 40 years triploidy was reported in less than 5% of cases (data not shown).

TABLE 2 FETAL RESULTS OF 24-CHROMOSOME MICROARRAY TESTING ON PRODUCTS OF CONCEPTION BY ESTIMATED GESTATIONAL AGE AT THE TIME OF PREGNANCY LOSS

Gestational age	Total number of samples	% with fetal results
<5 weeks	101/190	53.2
5 weeks – 5 weeks 6 days	1318/2071	63.6
6 weeks – 6 weeks 6 days	9236/11883	77.7
7 weeks – 7 weeks 6 days	9334/11345	82.3
8 weeks – 8 weeks 6 days	11,659/13,465	86.6
9 weeks – 9 weeks 6 days	6833/7756	88.1
10 weeks – 10 weeks 6 days	3435/3919	87.6
11 weeks – 11 weeks 6 days	2273/2480	91.7
12 weeks – 12 weeks 6 days	2116/2253	93.9
13 weeks – 19 weeks 6 days	8607/8809	97.7

Gestational age was the best estimate based on the last menstrual period or ultrasound scan result as reported by the clinician.

DISCUSSION

Management based on initial ASRM RPL evaluation with the addition of CCA of POC

The chromosomal status of the previously failed pregnancy is generally not available at the time of the initial consultation for RPL. Studies based on G-banding karyotype analysis suggest that the POC results from the first miscarriage (euploid or aneuploid) are associated with a similar genetic outcome in 65% of second miscarriages (*Ogasawara et al., 2000*). Using the CCA of POC revealed similar aneuploidy rates in pregnancy losses from natural conceptions (56.8%), losses from conceptions from IVF (53.6%), women with one prior loss (54.4%) and women with three or more losses (52.1%) (*Zhu et al., 2018*).

Aneuploidy in POC is extremely common, and its presence does not rule out the possibility of any of the known treatable RPL aetiologies, which should be evaluated. The addition of the POC cytogenetic

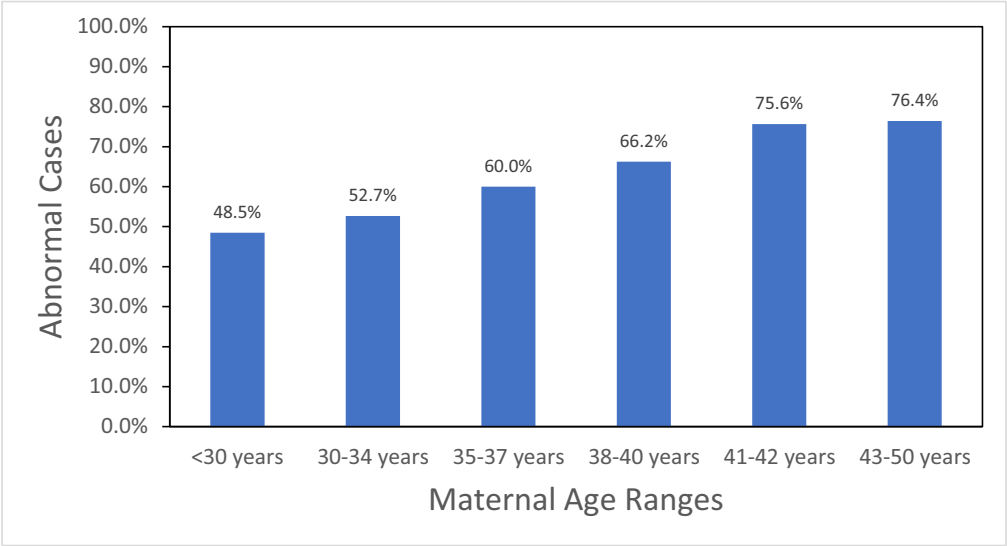


FIGURE 3 Rates of abnormal genetic results from products of conception (n = 54,912) according to maternal age ranges at the time of the pregnancy loss.

analysis to the standard RPL evaluation allows the clinician to direct management towards what would potentially be the most appropriate approach, without making assumptions regarding the chromosomal status of previous miscarriages (FIGURE 6). When the POC CCA reveals aneuploidy that is not an unbalanced Robertsonian translocation or inversion (about 55% of cases), the ASRM evaluation for RPL should be considered. (*Popescu et al., 2018*; *Zhu et al., 2018*).

Approximately 14% of all women with RPL who show POC aneuploidy on CCA will have at least one concomitant abnormal finding (explained) on the ASRM evaluation (*Popescu et al., 2018*) and should benefit from treatment in any subsequent pregnancy. In a patient who has been evaluated and treated for all the known aetiologies (FIGURE 6), no further testing is generally recommended, and expectant management for 6 months is appropriate.

For women with RPL who have a normal ASRM workup with POC aneuploidy as the sole identified cause of at least the last miscarriage (41% of all RPL cases), the authors optimistically counsel the couple about the chances of a future live-born child based on the maternal age and number of prior losses (*Lund et al., 2012*; *Perfetto et al., 2015*). IVF with or without PGT-A can be considered after 6 months of expectant management in couples with a normal RPL workup and documented aneuploidy of the miscarriage tissue. For couples with recurrent aneuploidy on CMA of the POC, additional counselling and a discussion about the potential benefits of IVF with or without PGT-A seems appropriate.

When the POC CCA reveals an unbalanced translocation or an inversion (about 3% of cases), a karyotype analysis of both parents should be performed. In cases where a balanced translocation is identified in one of the parents, the parents should receive genetic counselling. Expectant management versus IVF with PGT-SR should be considered and discussed in these cases (*Dahdouh et al. 2015a*; *Practice Committee of the American Society for Reproductive 2012*; *RPL et al. 2023*).

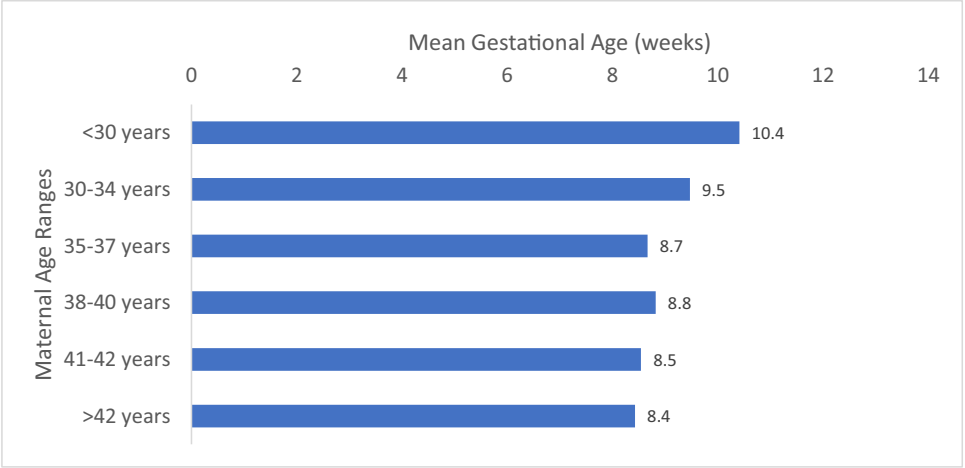


FIGURE 4 Mean gestational age at the time of the pregnancy loss according to the maternal age range.

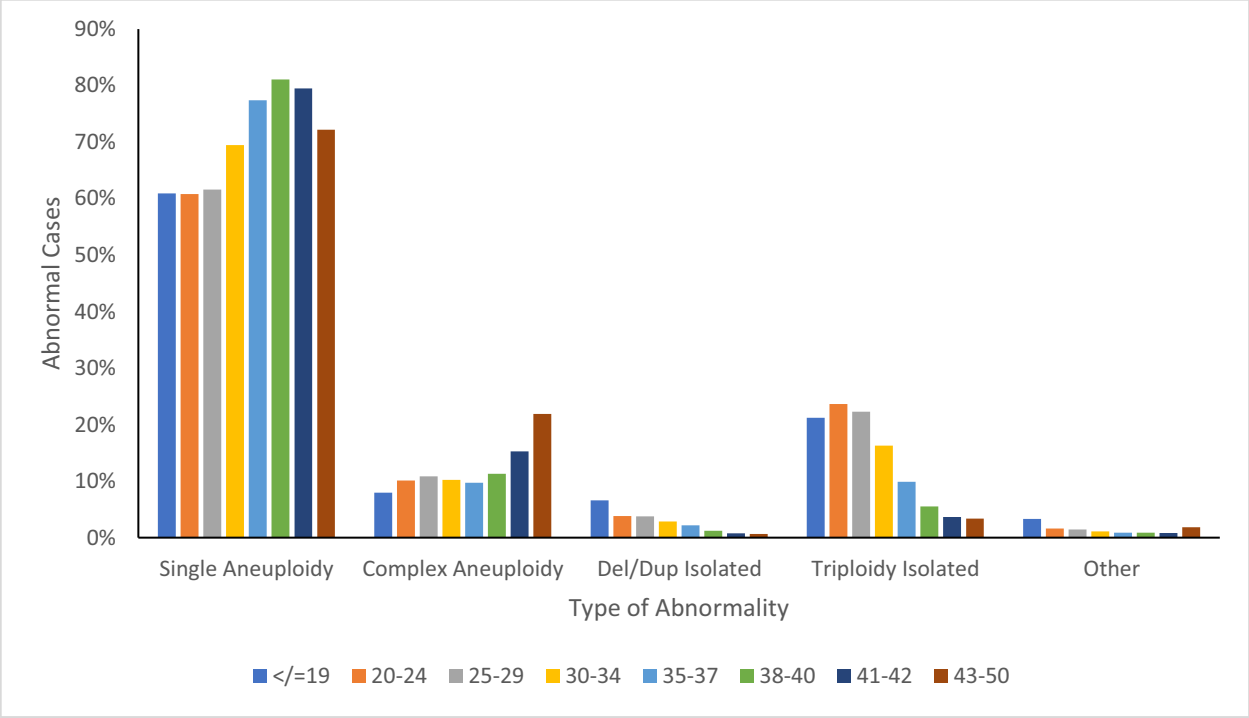


FIGURE 5 Types of genetic abnormality in products of conception according to maternal age at the time of the pregnancy loss. The colour key represents the maternal age ranges (years). Del, deletion; Dup, duplication.

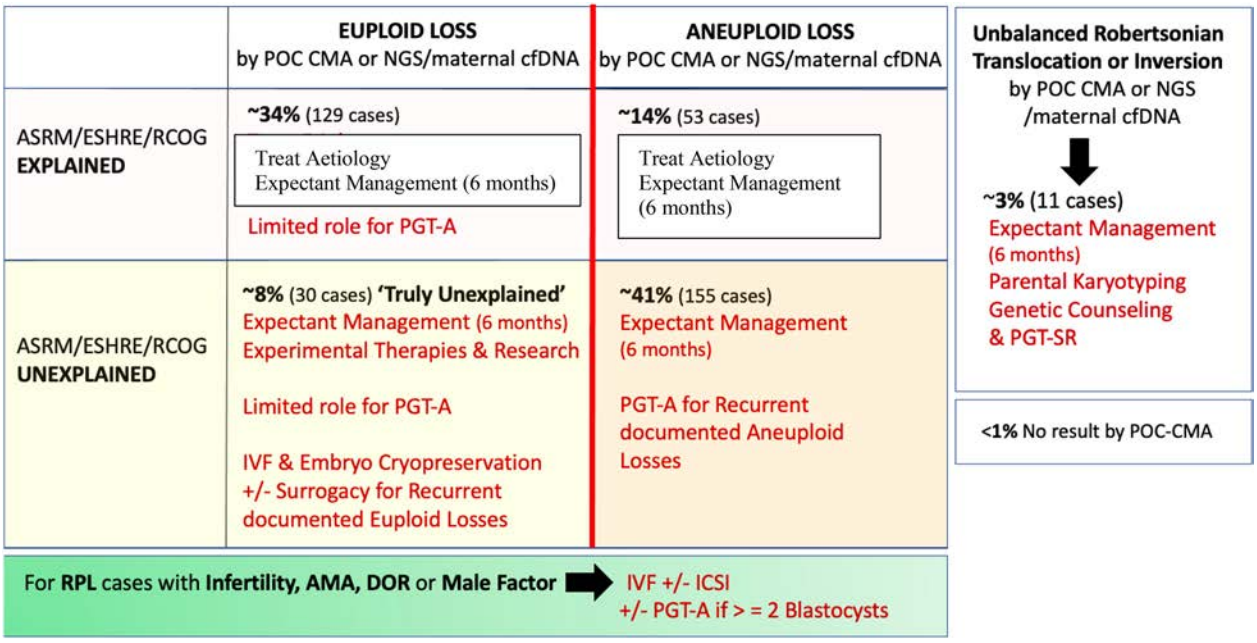


FIGURE 6 Proposed recommendations for the management of recurrent pregnancy loss (RPL), using 'explained versus unexplained' combined with 'euploid versus aneuploid' to determine candidates for expectant management, treatment or preimplantation genetic testing for aneuploidies (PGT-A) or structural rearrangements (PGT-SR), versus experimental therapy. The numbers are based on the of American Society for Reproductive Medicine (ASRM) evaluation and the results of 24-chromosomal microarray (CMA) analysis of products of conception (POC) from 378 prospectively recruited couples (Pappas and Kutteh, 2020). AMA, Advance Maternal Age; cfDNA, cell-free DNA; DOR, Decreased Ovarian Reserve; ESHRE, European Society of Human Reproduction and Embryology; ICSI, intracytoplasmic sperm injection; NGS, next-generation sequencing; RCOG, Royal College of Obstetricians and Gynaecologists.

When POC are euploid on cytogenetic analysis with an explained cause after the ASRM workup (about one-third of cases), the abnormal finding on the standard evaluation should be treated and the patient should be followed with expectant management. The majority of these women will have abnormal ASRM test results (*Popescu et al., 2018*). Overall, a euploid POC and an abnormal ASRM (explained) workup occurs in 34% of all cases (**FIGURE 6**). Given that at least one of the known and treatable aetiologies would have been addressed, no further testing is recommended at this point for those specific cases. Expectant management is appropriate, and the authors see a limited role for IVF with PGT-A in this group of patients with RPL. Exceptions to this management strategy include the recommendation for IVF with or without PGT-A in women with RPL and advanced maternal age or decreased ovarian reserve, and with male factor infertility.

RPL cases with an unexplained ASRM evaluation and euploid POC account for less than 10% of all patients (**FIGURE 6**). The authors suggest that this group of patients qualifies for a new definition of 'truly unexplained RPL' when incorporating systematic POC cytogenetic analysis as a routine part of the RPL evaluation (*Papas and Kutteh, 2021*). For these specific cases, additional experimental testing or entry into RCT can be considered if deemed appropriate by the managing medical team. Some clinicians suggest that IVF, embryo cryopreservation and frozen embryo transfer into a gestational carrier may be appropriate in this group of patients with recurrent euploid losses. Patient should be counselled that expectant management is also appropriate and that successful outcomes are frequent (*Lund et al., 2012*; *Perfetto et al., 2015*).

Maternal age is the major predictor of fetal aneuploidy

Most clinicians have been taught from their residency training that the majority of miscarriages are caused by genetic abnormalities in the pregnancy. The current authors' data on POC from 64,521 singleton pregnant losses demonstrate this to be true, with the frequency of losses attributable to chromosomal anomalies increasing with maternal age. In women under 30 years old, 48.5% of pregnancy losses were due to aneuploidy. In women over 42 years, aneuploidy in POC explained 76.4% of all miscarriages. Moreover, in women who were over 42 years old the pregnancy loss occur earlier, at on average 8.4 gestational weeks, than in women under 30 years, where losses occurred at 10.4 gestational weeks on average. Similarly, results from PGT-A cycles using CCA (with quantitative PCR) on more than 15,000 trophectoderm biopsies (*Franasiak et al., 2014*) and results from about 4000 blastocysts biopsied for SNP-based PGT-A (*Simon et al., 2018*) showed higher embryo aneuploidy rates with increasing maternal age, from about 27% in women less than 35 years old to more than 70–80% aneuploidy after the age of 42.

Role of PGT in treating women with RPL

A significantly higher rate of chromosomal abnormalities has been reported in the blastocysts of young women with idiopathic RPL compared with women with no sporadic miscarriage (*Liu et al., 2020*). In an intent-to-treat analysis, Murugappan and colleagues found similar LBR and clinical miscarriage rates in RPL patients whether expectant management or PGT-A was used (*Murugappan et al., 2016*). The interpretation of this study was limited by the predominant use of day 3 biopsies, which has been shown to lower pregnancy rates (*Mastenbroek et al., 2007*; *Mastenbroek et al., 2011*; *Treff et al. 2011*).

In one recent prospective cohort trial by Sato and co-workers PGT-A for patients with RPL due to embryonic aneuploidy was associated with a higher LBR per embryo transfer compared with regular IVF (52.8% versus 21.6%, $P = 0.028$), but not per patient (26.8% versus 21.1%, $P = 0.6$) (*Sato et al., 2019*). As expected, a high proportion of patients in this study did not have an available euploid embryo and hence did not reach the stage of embryo transfer (*Sato et al., 2019*). In a study of women younger than 38 years who had one aneuploid miscarriage, no improvement was found with the use of PGT-A in terms of LBR or the rate of miscarriage (*Zhou et al., 2021*). In contrast, Sanders and colleagues reported both a higher LBR per embryo transfer ($P < 0.05$) and per treatment cycle ($P < 0.05$) for maternal age groups older than 35 years (*Sanders et al., 2021*) and Sacchi and collaborators reported lower rates of pregnancy loss in PGT-A groups versus control groups (3.6% versus 22.6%; *Sacchi et al., 2019*). In a recent SART-CORS study published in 2021 by Bhatt and colleagues PGT-A was associated with an improved LBR in patients with RPL undergoing a frozen embryo transfer (*Bhatt et al., 2021*). The study was restricted to RPL cases undergoing IVF for an infertility diagnosis and having already frozen embryos, which may bias the results (*Dahdouh et al. 2021*). Currently, PGT-A use is not recommended by any society guidelines in either IVF patients with infertility or RPL patients (*Chan et al., 2021*).

Very recently, the proposed criteria considered necessary to perform successful PGT-A on RPL patients were updated and published by the current authors' group (*Dahdouh and Kutteh, 2021*). These include couples with a negative ASRM or ESHRE workup, a history of aneuploid pregnancy loss, and a female partner with a normal ovarian reserve producing a high number of oocytes and blastocysts available for biopsy. In addition, IVF is to be undertaken in centres with extensive experience with overall PGT-A practice (*Dahdouh et al., 2021*).

The current authors agree with recommendations in favour of expectant management in most cases of RPL explained by ASRM or ESHRE guidelines, especially in young patients. However, they believe there is a clear role for PGT-A in RPL, especially in cases with recurrent POC aneuploidy, as identified by one of the group's proposed algorithms (*Popescu et al., 2018*), as well as in cases unexplained by ASRM or ESHRE workups with one or more aneuploid losses identified by the latest proposed criteria by the group (*Papas and Kutteh, 2020*).

Impact of mosaic results on embryos tested by PGT-A

Despite its widespread use mainly in the USA and some Canadian provinces, PGT-A remains a controversial topic mainly because all of the RCT have comprised only good-prognosis patients (*Dahdouh et al., 2015c*; *Dahdouh et al., 2022*). Furthermore, with the introduction of highly sensitive platforms (i.e. NGS) into clinical practice, a classification with an intermediate copy number termed 'mosaic' can be reported. This comprised a mixture of euploid and aneuploid cells (20–80%) within the embryo biopsy sample. The ideal disposition and management of embryos with mosaic results remain a matter of debate, as many of these embryos gave rise to healthy live births with no identifiable congenital or chromosomal anomalies (*Leigh et al., 2022*) (**TABLE 3**). Large cohort studies on reproductive outcomes of 'mosaic'

TABLE 3 ESTIMATED PREGNANCY OUTCOMES BASED ON THE RESULTS OF PGT-A TESTING AFTER SINGLE-EMBRYO TRANSFER

PGT-A results	LBR for ongoing pregnancy (>20 weeks)	Miscarriage rate (%)
Euploid	69.1% with aCGH	2.6
	50% with NGS	9.9
	77.2% with NGS (cumulative LBR in 1 year)	8.7 (cumulative PL in 1 year)
Low-level mosaic (20–50%)	42–42.9% non-selection trial	11–12.7
	44.5%	5.1
High-level mosaic (50–80%)	36%	30.7

Data from four studies were used (*Dahdouh, 2021*; *Dahdouh et al., 2022*; *Munne et al., 2019*; *Yan et al., 2021*).

aCGH, array comparative genomic hybridization; LBR, live birth rate; NGS, next-generation sequencing; PGT-A, preimplantation genetic testing for aneuploidies; PL, pregnancy loss.

embryos have reported lower LBR and higher miscarriage rates compared with euploid embryo results (*Viotti et al., 2021*). Hence, an embryo with a mosaic result might be eligible for transfer especially in those patients with no euploid embryo available, but only after genetic counselling (*Zwingerman and Langlois, 2020*). At the present time, many societies' guidelines recommend amniocentesis for any established pregnancy arising from an embryo with a mosaic result (*Zwingerman and Langlois, 2020*), as chromosomally abnormal fetuses have recently been reported following mosaic embryo transfers (*Greco et al., 2023*; *Huang et al., 2021*).

Not all genetic testing modalities are equal, and each has limitations, reducing the accuracy in detecting an abnormal pregnancy. Karyotype banding analysis and fluorescence in-situ hybridization technologies need viable cells and are plagued by cell culture failure as well as an inability to detect MCC or to determine the parental origin of any aneuploidy (*Levy et al., 2014*). Some investigators suggest that it is practical to perform genetic testing for all miscarriages (*Popescu et al., 2018*; *Xie et al., 2016*). Nevertheless, CCA of POC can be accomplished by different genetic platforms (i.e. comparative genomic hybridization or SNP microarrays, and NGS). SNP microarray has the advantage over array comparative genomic hybridization in detecting triploidy (molar pregnancies), present in 3% of all miscarriages (*Maisenbacher et al., 2019*). Additionally, NGS can diagnose mosaic losses, with trisomy 21 and monosomy XO being the most common true mosaic fetuses after 15 weeks' gestation (*Huang et al. 2021*).

Strengths and limitations

The strengths of the current study include the selection of a large cohort of over 65,000 patients who had commercial POC testing undertaken at a single laboratory. Female age and gestational age at the time of tissue collection were recorded by the clinical staff at the time of the loss. The limitations of the study include the inability to know when the pregnancy failed in relation to when the miscarriage sample was collected. For example, a patient with a gestational age of 10 weeks who had a pregnancy that failed at 7 weeks might report to her obstetrician at 10 weeks when she complains of spotting. Thus, in most cases, the gestational age that was reported was the gestational age based on the last menstrual period. In addition, fetal results were only available in 84.0% of cases, with lower percentages of fetal test results being available at lower gestational ages.

Preliminary studies have suggested that maternal plasma genome-wide cell-free DNA (cfDNA) testing at the time of the pregnancy loss may be a useful adjunct in cases where POC tissue is not obtained or not available. Furthermore, SNP microarray analysis allows for the identification of other genetic abnormalities that could explain the pregnancy loss and/or carry recurrence risks for the patient. As the costs of genetic testing continue to decrease, SNP microarray analysis may become a more widely available diagnostic testing option for pregnancy loss.

Future perspectives

NGS analysis of maternal cfDNA is a technology currently used for non-invasive prenatal testing that involves detecting embryonal chromosomal abnormalities in fetal cfDNA originating from shed and ruptured placental cells present in maternal blood. It has recently been proposed as a replacement for POC genetic analysis in RPL (*Colley et al., 2020*; *Yaron et al., 2020*). cfDNA technology offers the advantage of not having to collect a POC sample and of much more practical sampling. Some patients miscarry before a sample can be obtained, and many miscarriages are currently managed medically; therefore this technology might have a promising role.

However, larger data sets are needed to validate the use of cfDNA for RPL patients. In a large recently published prospective cohort study, 667 women with pregnancy loss were recruited to evaluate fetal cfDNA performance (*Schlaikjaer Hartwig et al. 2023*). This promising non-invasive test showed a sensitivity for aneuploidy detection of 85% (95% CI 79–90%) and a specificity of 93% (95% CI 88–96%) compared with genetic testing (by NGS) of the pregnancy tissue (*Schlaikjaer Hartwig et al., 2023*). The current authors believe that the addition of POC genetic testing (or in the future the NGS analysis of maternal cfDNA) to the standard workup of RPL cases can be valuable, given that these results can direct patient care appropriately towards either expectant management or IVF with or without PGT-A, potentially improving and personalizing the management of RPL.

Furthermore, the recent identification of embryonic cfDNA in spent blastocyst culture media has opened a new era for the development of non-invasive embryo aneuploidy screening, named niPGT-A (*Cinnioglu et al., 2023*; *Rubio et al., 2020*). This can avoid the impact of invasive embryo biopsy and might overcome technical challenges and decrease cost, potentially increasing accessibility for a wider patient population (*Rubio et al., 2020*). niPGT-A has shown equivalent concordance rates to blastocyst-stage biopsy, but still needs additional development and research in order to be adopted as a routine selection tool in IVF cycles (*Cinnioglu et al., 2023*; *Xu et al., 2023*). The results from an ongoing large multicentre double-blind RCT comparing the outcomes from embryos selected using niPGT-A with embryos selected using standard morphology criteria are urgently awaited and will clarify further its promising role (*Huang et al., 2022*).

CONCLUSIONS

CCA of the POC combined with a standard workup for RPL can provide a possible or proven explanation for the loss in over 90% of couples. Categorizing the loss as euploid versus aneuploid, in addition to classifying the standard evaluation as explained versus unexplained, can help to identify candidates suitable for expectant management versus PGT-A.

DATA AVAILABILITY

The data that has been used is confidential.

ACKNOWLEDGEMENTS

This article is dedicated to the memory of Dr Ralph S. Papas who made numerous contributions to the field before his recent passing. Dr Papas contributed significantly to the data collection and preparation of the manuscript. He will be remembered for his enthusiasm for simplifying complex concepts in the field of recurrent pregnancy loss. The authors also acknowledge the assistance of Katherine Howard, MS, CGC, Clinical Data Manager at Natera, Inc. for updating the products of conception test results for this report, and the contributions of Katrina Merrion, MS, CGC to the editing and data review, and of Zach Demko to the critical review of the final manuscript.

AUTHOR CONTRIBUTIONS

All the authors have made substantial contributions to all of the following: (i) the conception and design of the study, or the acquisition of data, or the analysis and interpretation of data; (ii) drafting the article or revising it critically for important intellectual content; and (iii) the final approval of the version to be submitted.

REFERENCES

ACOG. 2023. Repeated Miscarriages-Frequently Asked Questions. https://www.acog.org/womens-health/faqs/repeated-miscarriages.

Bhatt, S.J., Marchetto, N.M., Roy, J., Morelli, S.S., McGovern, P.G., 2021. Pregnancy outcomes following in vitro fertilization frozen embryo transfer (IVF-FET) with or without preimplantation genetic testing for aneuploidy (PGT-A) in women with recurrent pregnancy loss (RPL): a SART-CORS study. Hum Reprod 36 (8), 2339–2344. https://doi.org/10.1093/humrep/deab117.

Brezina PR, K.W., 2013. A new algorithm for the evaluation of recurrent pregnancy loss. Clin Reprod Med Surg 2, 197–208.

Chan, C., Ryu, M., Zwingerman, R., 2021. Preimplantation genetic testing for aneuploidy: A Canadian Fertility and Andrology Society Guideline. Reprod Biomed Online 42 (1), 105–116. https://doi.org/10.1016/j.rbmo.2020.10.020.

Cinnioglu, C., Glessner, H., Jordan, A., Bunshaft, S., 2023. A Systematic Review of Non-invasive Preimplantation Genetic Testing for Aneuploidy. Fertil Steril. https://doi.org/10.1016/j.fertnstert.2023.06.013.

Colley, E., Devall, A.J., Williams, H., Hamilton, S., Smith, P., Morgan, N.V., Quenby, S., Coomarasamy, A., Allen, S., 2020. Cell-Free DNA in the Investigation of Miscarriage. J Clin Med 9 (11). https://doi.org/10.3390/jcm9113428.

Committee EPCS, Carvalho, F., Coonen, E., Goossens, V., Kokkali, G., Rubio, C., Meijer-Hoogeveen, M., Moutou, C., Vermeulen, N., De Rycke, M., 2020. ESHRE PGT Consortium good practice recommendations for the organisation of PGT. Hum Reprod Open 2020 (3), hoaa021. https://doi.org/10.1093/hropen/hoaa021.

Dahdouh, E.M., 2021. Preimplantation Genetic Testing for Aneuploidy: A Review of the Evidence. Obstet Gynecol 137 (3), 528–534. https://doi.org/10.1097/AOG.0000000000004295.

Dahdouh, E.M., Balayla, J., Audibert, F., Genetics, C., Wilson, R.D., Audibert, F., Brock, J.A., Campagnolo, C., Carroll, J., Chong, K., Gagnon, A., Johnson, J.A., MacDonald, W., Okun, N., Pastuck, M., Vallee-Pouliot, K., 2015a. RETIRED: Technical Update: Preimplantation Genetic Diagnosis and Screening. J Obstet Gynaecol Can 37 (5), 451–463. https://doi.org/10.1016/s1701-2163(15)30261-9.

Dahdouh, E.M., Balayla, J., Garcia-Velasco, J.A., 2015b. Comprehensive chromosome screening improves embryo selection: a meta-analysis. Fertil Steril 104 (6), 1503–1512. https://doi.org/10.1016/j.fertnstert.2015.08.038.

Dahdouh, E.M., Balayla, J., Garcia-Velasco, J.A., 2015c. Impact of blastocyst biopsy and comprehensive chromosome screening technology on preimplantation genetic screening: a systematic review of randomized controlled trials. Reprod Biomed Online 30 (3), 281–289. https://doi.org/10.1016/j.rbmo.2014.11.015.

Dahdouh, E.M., Balayla, J., Garcia-Velasco, J.A., Kutteh, W.H., 2021. PGT-A for recurrent pregnancy loss: evidence is growing but the issue is not resolved. Hum Reprod 36 (10), 2805–2806. https://doi.org/10.1093/humrep/deab194.

Dahdouh, E.M., Kutteh, W.H., 2021. Genetic testing of products of conception in recurrent pregnancy loss evaluation. Reprod Biomed Online 43 (1), 120–126. https://doi.org/10.1016/j.rbmo.2021.03.015.

Dahdouh, E.M., Mourad, A.M., Balayla, J., Sylvestre, C., Brezina, P.R., Kutteh, W.H., Picchetta, L., Capalbo, A., Garcia-Velasco, J.A., 2022. Update on preimplantation genetic testing for aneuploidy and outcomes of embryos with mosaic results. Minerva Obstet Gynecol. https://doi.org/10.23736/S2724-606X.22.05166-1.

Dimitriadis, E., Menkhorst, E., Saito, S., Kutteh, W.H., Brosens, J.J., 2020. Recurrent pregnancy loss. Nat Rev Dis Primers 6 (1), 98. https://doi.org/10.1038/s41572-020-00228-z.

Franasiak, J.M., Forman, E.J., Hong, K.H., Werner, M.D., Upham, K.M., Treff, N.R., Scott, Jr., R.T., 2014. The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertil Steril 101 (3), 656–663.e651. https://doi.org/10.1016/j.fertnstert.2013.11.004.

Greco, E., Yakovlev, P., Kornilov, N., Vyatkina, S., Bogdanova, D., Ermakova, M., Tarasova, Y., Tikhonov, A., Pendina, A., Biricik, A., Sessa, M.T., Listorti, I., Ronsini, C., Greco, P.F., Victor, A., Barnes, F., Zouves, C., Spinella, F., Viotti, M., 2023. Two clinical case reports of embryonic mosaicism identified with PGT-A persisting during pregnancy as true fetal mosaicism. Hum Reprod 38 (2), 315–323. https://doi.org/10.1093/humrep/deac263.

Huang, C., Jiang, W., Zhu, Y., Li, H., Lu, J., Yan, J., Chen, Z.J., 2019. Pregnancy outcomes of reciprocal translocation carriers with two or more unfavorable pregnancy histories: before and after preimplantation genetic testing. J Assist Reprod Genet 36 (11), 2325–2331. https://doi.org/10.1007/s10815-019-01585-9.

Huang, J., Rong, L., Zeng, L., Hu, L., Shi, J., Cai, L., Yao, B., Wang, X.X., Xu, Y., Yao, Y., Wang, Y., Zhao, J., Guan, Y., Qian, W., Hao, G., Lu, S., Liu, P., Qiao, J., 2022. Embryo selection through non-invasive preimplantation genetic testing with cell-free DNA in spent culture media: a protocol for a multicentre, double-blind, randomised controlled trial. BMJ Open 12 (7), e057254. https://doi.org/10.1136/bmjopen-2021-057254.

Huang, K.L., Tsai, C.C., Cheng, H.H., Huang, Y.J., Lai, Y.J., Wu, C.H., Hsiao, P.Y., Hsu, T.Y., Lan, K.C., 2021. Whether to transfer mosaic embryos: a cytogenetic view of true mosaicism by amniocentesis. Reprod Biomed Online 43 (1), 33–43. https://doi.org/10.1016/j.rbmo.2021.03.003.

Jaslow, C.R., Carney, J.L., Kutteh, W.H., 2010. Diagnostic factors identified in 1020 women with two versus three or more recurrent pregnancy losses. Fertil Steril 93 (4), 1234–1243. https://doi.org/10.1016/j.fertnstert.2009.01.166.

Jaslow, C.R., Kutteh, W.H., 2013. Effect of prior birth and miscarriage frequency on the prevalence of acquired and congenital uterine anomalies in women with recurrent miscarriage: a cross-sectional study. Fertil Steril 99 (7), 1916–1922.e1911. https://doi.org/10.1016/j.fertnstert.2013.01.152.

Johnson, D.S., Gemelos, G., Baner, J., Ryan, A., Cinnioglu, C., Banjevic, M., Ross, R., Alper, M., Barrett, B., Frederick, J., Potter, D., Behr, B., Rabinowitz, M., 2010. Preclinical validation of a microarray method for full molecular karyotyping of blastomeres in a 24-h protocol. Hum Reprod 25 (4), 1066–1075. https://doi.org/10.1093/humrep/dep452.

Kutteh, W.H., 2014. Antiphospholipid antibody syndrome and reproduction. Curr Opin Obstet Gynecol 26 (4), 260–265. https://doi.org/10.1097/GCO.0000000000000086.

Leigh, D., Cram, D.S., Rechitsky, S., Handyside, A., Wells, D., Munne, S., Kahraman, S., Grifo, J., Katz-Jaffe, M., Rubio, C., Viotti, M., Forman, E., Xu, K., Gordon, T., Madjunkova, S., Qiao, J., Chen, Z.J., Harton, G., Gianaroli, L., Simon, C., Scott, R., Simpson, J.L., Kuliev, A., 2022. PGDIS position statement on the transfer of mosaic embryos 2021. Reprod Biomed Online 45 (1), 19–25. https://doi.org/10.1016/j.rbmo.2022.03.013.

Levy, B., Sigurjonsson, S., Pettersen, B., Maisenbacher, M.K., Hall, M.P., Demko, Z., Lathi, R.B., Tao, R., Aggarwal, V., Rabinowitz, M., 2014. Genomic imbalance in products of conception: single-nucleotide polymorphism chromosomal microarray analysis. Obstet Gynecol 124 (2 Pt 1), 202–209. https://doi.org/10.1097/AOG.0000000000000325.

Liu, X.Y., Fan, Q., Wang, J., Li, R., Xu, Y., Guo, J., Wang, Y.Z., Zeng, Y.H., Ding, C.H., Cai, B., Zhou, C.Q., Xu, Y.W., 2020. Higher chromosomal abnormality rate in blastocysts from young patients with idiopathic recurrent pregnancy loss. Fertil Steril 113 (4), 853–864. https://doi.org/10.1016/j.fertnstert.2019.11.016.

Lund, M., Kamper-Jorgensen, M., Nielsen, H.S., Lidegaard, O., Andersen, A.M., Christiansen, O.B., 2012. Prognosis for live birth in women with recurrent miscarriage: what is the best measure of success? Obstet Gynecol 119 (1), 37–43. https://doi.org/10.1097/AOG.0b013e31823c0413.

Maisenbacher, M., Kutteh, WH., 2020. Single nucleotide polymorphism (SNP) array analysis of 63,277 products of conception (POC) samples: a 10-year laboratory experience. Fertil Steril 114, E47.

Maisenbacher, M.K., Merrion, K., Kutteh, W.H., 2019. Single-nucleotide polymorphism microarray detects molar pregnancies in 3% of miscarriages. Fertil Steril 112 (4), 700–706. https://doi.org/10.1016/j.fertnstert.2019.06.015.

Mastenbroek, S., Twisk, M., van der Veen, F., Repping, S., 2011. Preimplantation genetic screening: a systematic review and meta-analysis of RCTs. Hum Reprod Update 17 (4), 454–466. https://doi.org/10.1093/humupd/dmr003.

Mastenbroek, S., Twisk, M., van Echten-Arends, J., Sikkema-Raddatz, B., Korevaar, J.C., Verhoeve, H.R., Vogel, N.E., Arts, E.G., de Vries, J.W., Bossuyt, P.M., Buys, C.H., Heineman, M.J., Repping, S., van der Veen, F., 2007. In vitro fertilization with preimplantation genetic screening. N Engl J Med 357 (1), 9–17. https://doi.org/10.1056/NEJMoa067744.

Munne, S., Kaplan, B., Frattarelli, J.L., Child, T., Nakhuda, G., Shamma, F.N., Silverberg, K., Kalista, T., Handyside, A.H., Katz-Jaffe, M., Wells, D., Gordon, T., Stock-Myer, S., Willman, S., Group SS, 2019. Preimplantation genetic testing for aneuploidy versus morphology as selection criteria for single frozen-thawed embryo transfer in good-prognosis patients: a multicenter randomized clinical trial. Fertil Steril 112 (6), 1071–1079.e1077. https://doi.org/10.1016/j.fertnstert.2019.07.1346.

Murugappan, G., Shahine, L.K., Perfetto, C.O., Hickok, L.R., Lathi, R.B., 2016. Intent to treat analysis of in vitro fertilization and preimplantation genetic screening versus expectant management in patients with recurrent pregnancy loss. Hum Reprod 31 (8), 1668–1674. https://doi.org/10.1093/humrep/dew135.

Ogasawara, M., Aoki, K., Okada, S., Suzumori, K., 2000. Embryonic karyotype of abortuses in relation to the number of previous miscarriages. Fertil Steril 73 (2), 300–304. https://doi.org/10.1016/s0015-0282(99)00495-1.

Papas, R.S., Kutteh, W.H., 2020. A new algorithm for the evaluation of recurrent pregnancy loss redefining unexplained miscarriage: review of current guidelines. Curr Opin Obstet Gynecol 32 (5), 371–379. https://doi.org/10.1097/GCO.0000000000000647.

Papas, R.S., Kutteh, W.H., 2021. Genetic Testing for Aneuploidy in Patients Who Have Had Multiple Miscarriages: A Review of Current Literature. Appl Clin Genet 14, 321–329. https://doi.org/10.2147/TACG.S320778.

Perfetto, C.O., Murugappan, G., Lathi, R.B., 2015. Time to next pregnancy in spontaneous pregnancies versus treatment cycles in fertile patients with recurrent pregnancy loss. Fertil Res Pract 1, 5. https://doi.org/10.1186/2054-7099-1-5.

Popescu, F., Jaslow, C.R., Kutteh, W.H., 2018. Recurrent pregnancy loss evaluation combined with 24-chromosome microarray of miscarriage tissue provides a probable or definite cause of pregnancy loss in over 90% of patients. Hum Reprod 33 (4), 579–587. https://doi.org/10.1093/humrep/dey021.

Practice Committee of the American Society for Reproductive M, 2012. Evaluation and treatment of recurrent pregnancy loss: a committee opinion. Fertil Steril 98 (5), 1103–1111. https://doi.org/10.1016/j.fertnstert.2012.06.048.

Preimplantation Genetic Testing: ACOG Committee Opinion, Number 799, 2020. Obstet Gynecol 135 (3), e133–e137. https://doi.org/10.1097/AOG.0000000000003714.

RPL EGGo, Bender Atik, R., Christiansen, O.B., Elson, J., Kolte, A.M., Lewis, S., Middeldorp, S., McHeik, S., Peramo, B., Quenby, S., Nielsen, H.S., van der Hoorn, M.L., Vermeulen, N., Goddijn, M., 2023. ESHRE guideline: recurrent pregnancy loss: an update in 2022. Hum Reprod Open 2023 (1), hoad002. https://doi.org/10.1093/hropen/hoad002.

Rubio, C., Navarro-Sanchez, L., Garcia-Pascual, C.M., Ocali, O., Cimadomo, D., Venier, W., Barroso, G., Kopcow, L., Bahceci, M., Kulmann, M.I.R., Lopez, L., De la Fuente, E., Navarro, R., Valbuena, D., Sakkas, D., Rienzi, L., Simon, C., 2020. Multicenter prospective study of concordance between embryonic cell-free DNA and trophectoderm biopsies from 1301 human blastocysts. Am J Obstet Gynecol 223 (5). https://doi.org/10.1016/j.ajog.2020.04.035 751 e751-751 e713.

Sacchi, L., Albani, E., Cesana, A., Smeraldi, A., Parini, V., Fabiani, M., Poli, M., Capalbo, A., Levi-Setti, P.E., 2019. Preimplantation Genetic Testing for Aneuploidy Improves Clinical, Gestational, and Neonatal Outcomes in Advanced Maternal Age Patients Without Compromising Cumulative Live-Birth Rate. J Assist Reprod Genet 36 (12), 2493–2504. https://doi.org/10.1007/s10815-019-01609-4.

Sanders, K.D., Silvestri, G., Gordon, T., Griffin, D.K., 2021. Analysis of IVF live birth outcomes with and without preimplantation genetic testing for aneuploidy (PGT-A): UK Human Fertilisation and Embryology Authority data collection 2016-2018. J Assist Reprod Genet 38 (12), 3277–3285. https://doi.org/10.1007/s10815-021-02349-0.

Sato, T., Sugiura-Ogasawara, M., Ozawa, F., Yamamoto, T., Kato, T., Kurahashi, H., Kuroda, T., Aoyama, N., Kato, K., Kobayashi, R., Fukuda, A., Utsunomiya, T., Kuwahara, A., Saito, H., Takeshita, T., Irahara, M., 2019. Preimplantation genetic testing for aneuploidy: a comparison of live birth rates in patients with recurrent pregnancy loss due to embryonic aneuploidy or recurrent implantation failure. Hum Reprod 34 (12), 2340–2348. https://doi.org/10.1093/humrep/dez229.

Schlaikjaer Hartwig, T., Ambye, L., Gruhn, J.R., Petersen, J.F., Wronding, T., Amato, L., Chi-Ho Chan, A., Ji, B., Bro-Jorgensen, M.H., Werge, L., Petersen, M., Brinkmann, C., Petersen, J.B., Duno, M., Bache, I., Herrgard, M.J., Jorgensen, F.S., Hoffmann, E.R., Nielsen, H.S., consortium, C., 2023. Cell-free fetal DNA for genetic evaluation in Copenhagen Pregnancy Loss Study (COPL): a prospective cohort study. Lancet 401 (10378), 762–771. https://doi.org/10.1016/S0140-6736(22)02610-1.

Simon, A.L., Kiehl, M., Fischer, E., Proctor, J.G., Bush, M.R., Givens, C., Rabinowitz, M., Demko, Z.P., 2018. Pregnancy outcomes from more than 1,800 in vitro fertilization cycles with the use of 24-chromosome single-nucleotide polymorphism-based preimplantation genetic testing for aneuploidy. Fertil Steril 110 (1), 113–121. https://doi.org/10.1016/j.fertnstert.2018.03.026.

Treff, N.R., Northrop, L.E., Kasabwala, K., Su, J., Levy, B., Scott, Jr., R.T., 2011. Single nucleotide polymorphism microarray-based concurrent screening of 24-chromosome aneuploidy and unbalanced translocations in preimplantation human embryos. Fertil Steril 95 (5), 1606–1612. e1601-1602. https://doi.org/10.1016/j.fertnstert.2010.11.004.

Viotti, M., 2020. Preimplantation Genetic Testing for Chromosomal Abnormalities: Aneuploidy, Mosaicism, and Structural Rearrangements. Genes (Basel) 11 (6). https://doi.org/10.3390/genes11060602.

Viotti, M., Victor, A.R., Barnes, F.L., Zouves, C.G., Besser, A.G., Grifo, J.A., Cheng, E.H., Lee, M.S., Horcajadas, J.A., Corti, L., Fiorentino, F., Spinella, F., Minasi, M.G., Greco, E., Munne, S., 2021. Using outcome data from one thousand mosaic embryo transfers to formulate an embryo ranking system for clinical use. Fertil Steril 115 (5), 1212–1224. https://doi.org/10.1016/j.fertnstert.2020.11.041.

Xie, Y., Pei, X., Dong, Y., Wu, H., Wu, J., Shi, H., Zhuang, X., Sun, X., He, J., 2016. Single nucleotide polymorphism-based microarray analysis for the diagnosis of hydatidiform moles. Mol Med Rep 14 (1), 137–144. https://doi.org/10.3892/mmr.2016.5211.

Xu, C.L., Wei, Y.Q., Tan, Q.Y., Huang, Y., Wu, J.J., Li, C.Y., Ma, Y.F., Zhou, L., Liang, B., Kong, L.Y., Xu, R.X., Wang, Y.Y., 2023. Concordance of PGT for aneuploidies between blastocyst biopsies and spent blastocyst culture medium. Reprod Biomed Online 46 (3), 483–490. https://doi.org/10.1016/j.rbmo.2022.10.001.

Yan, J., Qin, Y., Zhao, H., Sun, Y., Gong, F., Li, R., Sun, X., Ling, X., Li, H., Hao, C., Tan, J., Yang, J., Zhu, Y., Liu, F., Chen, D., Wei, D., Lu, J., Ni, T., Zhou, W., Wu, K., Gao, Y., Shi, Y., Lu, Y., Zhang, T., Wu, W., Ma, X., Ma, H., Fu, J., Zhang, J., Meng, Q., Zhang, H., Legro, R.S., Chen, Z.J., 2021. Live Birth with or without Preimplantation Genetic Testing for Aneuploidy. N Engl J Med 385 (22), 2047–2058. https://doi.org/10.1056/NEJMoa2103613.

Yaron, Y., Pauta, M., Badenas, C., Soler, A., Borobio, V., Illanes, C., Paz, Y.M.F., Martinez-Portilla, R., Borrell, A., 2020. Maternal plasma genome-wide cell-free DNA can detect fetal aneuploidy in early and recurrent pregnancy loss and can be used to direct further workup. Hum Reprod 35 (5), 1222–1229. https://doi.org/10.1093/humrep/deaa073.

Zhou, T., Zhu, Y., Zhang, J., Li, H., Jiang, W., Zhang, Q., Lu, J., Yan, J., Chen, Z.J., 2021. Effects of PGT-A on Pregnancy Outcomes for Young Women Having One Previous Miscarriage with Genetically Abnormal Products of Conception. Reprod Sci 28 (11), 3265–3271. https://doi.org/10.1007/s43032-021-00542-1.

Zhu, X., Li, J., Zhu, Y., Wang, W., Wu, X., Yang, Y., Gu, L., Gu, Y., Hu, Y., 2018. Application of chromosomal microarray analysis in products of miscarriage. Mol Cytogenet 11, 44. https://doi.org/10.1186/s13039-018-0396-y.

Zwingerman, R., Langlois, S., 2020. Committee Opinion No. 406: Prenatal Testing After IVF With Preimplantation Genetic Testing for Aneuploidy. J Obstet Gynaecol Can 42 (11), 1437–1443.e1431. https://doi.org/10.1016/j.jogc.2019.11.069.

Received 27 September 2023; received in revised form 11 November 2023; accepted 15 November 2023.

RBMO

ELSEVIER

ARTICLE

Perceptions of reproductive healthcare providers regarding their involvement in offering expanded carrier screening in fertility clinics: a qualitative study

BIOGRAPHY

As a researcher at Amsterdam University Medical Centre (the Netherlands), David Klein explored the ethics of personalized medicine and reproductive health care. He participated in a consortium on ethical and legal issues of personalized medicine, assessing pathways for gene editing therapies, and conducted interviews on reproductive carrier screening. He is committed to responsible healthcare innovation.

David Klein[1], Ivy van Dijke[1,2], Irene M. van Langen[3], Wybo Dondorp[4], Phillis Lakeman[5], Lidewij Henneman[1,*], Martina C. Cornel[1]

KEY MESSAGE

While agreeing that the field of medically assisted reproduction provides a unique opportunity to offer expanded genetic carrier screening, reproductive healthcare workers in the Netherlands currently feel a lack of capability and limited motivation to offer this screening at their clinics. The perspectives of professionals are important for responsible implementation in fertility clinics.

ABSTRACT

Research question: What are the main arguments of reproductive healthcare providers in favour or against their involvement in offering expanded carrier screening (ECS) for recessive disorders at fertility clinics in the Netherlands?

Design: Semi-structured interview study with 20 reproductive healthcare providers between May 2020 and January 2021. Participants included 11 gynaecologists, seven fertility doctors, one nurse practitioner and one clinical embryologist, recruited from academic medical centres ($n = 13$), peripheral facilities associated with academic centres ($n = 4$), and independent fertility treatment centres ($n = 3$) in the Netherlands. An interview guide was developed, and thematic content analysis was performed using ATLAS.ti software.

Results: Arguments of reproductive healthcare providers in favour of their potential involvement in offering ECS included: (i) opportunities offered by the setting; (ii) motivation to assist in reproduction and prevent suffering; and (iii) to counter unwanted commercialization offers. Arguments against involvement included: (i) lack of knowledge and familiarity with offering ECS; (ii) insufficient staff and resources, and potential high costs for clinics and/or couples; (iii) the emotional impact it may have on couples; (iv) perceived complexity of counselling and expected elongation of waiting lists; and (v) expected low impact on

[1] Department of Human Genetics, Amsterdam Reproduction and Development Research Institute, Amsterdam University Medical Centre, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[2] Centre for Reproductive Medicine, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, the Netherlands
[3] Department of Genetics, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands
[4] Department of Health, Ethics and Society, Care and Public Health Research Institute, Maastricht University Medical Centre, Maastricht University, Maastricht, the Netherlands
[5] Department of Human Genetics, Amsterdam Reproduction and Development Research Institute, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, the Netherlands

KEY WORDS
Expanded carrier screening
Fertility clinics
Gynaecology
Attitude
Interviews

*Corresponding author. E-mail address: l.henneman@amsterdamumc.nl (L. Henneman). https://doi.org/10.1016/j.rbmo.2024.103857 1472-6483/
Declaration: All authors are affiliated to academic hospitals that offer expanded carrier screening on a non-commercial basis.

reducing the burden of diseases. Participants felt that more evidence and research on the costs–benefits, implications and demand are needed prior to their involvement.

Conclusion: While agreeing that the field of medically assisted reproduction provides a unique opportunity to offer ECS, reproductive healthcare workers feel a lack of capability and limited motivation to offer ECS to all or a selection of couples at their fertility clinics.

INTRODUCTION

Carrier screening aims to provide carrier couples with insight into their increased chance of having a child affected by an autosomal- or X-linked-recessive disorder, thereby enhancing reproductive autonomy. Until recently, carrier screening initiatives were mainly ancestry based, performed for one or a few disorders (*Delatycki et al., 2019*; *Gregg and Edwards, 2018*; *Henneman et al., 2016*). Due to technological advancements and continuous cost reductions, all prospective parents can now be screened for hundreds of autosomal- and X-linked-recessive disorders with one sample, which allows population-based implementation without limiting screening to specific groups (*Henneman et al., 2016*). This is commonly referred to as 'expanded carrier screening' (ECS, with 'expanded' referring to more conditions) or 'expanded universal carrier screening' (universal or population based as opposed to ancestry based), which is ideally offered to couples before pregnancy (*de Wert et al., 2021*; *van der Hout et al., 2019*). Based on data from European populations, an estimated one in 100–125 couples are at risk of conceiving an affected child (*Fridman et al., 2021*).

The manner in which (expanded) carrier screening is offered varies across the globe. For instance, in many European countries, carrier screening for recessive disorders is generally not part of national policy (*Delatycki et al., 2019*). Also, numerous initiatives have shown a low level of awareness and generally low uptake among the general population, despite the considered value of ECS for prospective parents (*Van Steijvoort et al., 2020*). Few couples inform their healthcare providers about their pregnancy plans prior to conception, making it challenging to reach all couples before conception (*Goossens et al., 2018*; *Rowe and Wright, 2020*).

In Australia, an ECS pilot was performed in around 10,000 couples as part of 'Mackenzie's Mission', a government-funded part of the Australian Genomics Health Futures Mission (*Kirk et al., 2021*). In the USA, carrier screening recommendations by the American College of Obstetrics and Gynecology (ACOG) are still limited to carrier screening for cystic fibrosis, spinal muscular atrophy and haemoglobinopathies (*ACOG, 2017*). However, ECS tests are increasingly offered by commercial providers, including fertility clinics (*Chokoshvili et al., 2017*; *Delatycki et al., 2019*). Most recently, the American College of Medical Genetics and Genomics (ACMG) recommended widening the scope of (what they propose to no longer refer to as 'expanded') carrier screening to conditions with a carrier frequency of one in ≥200 (*Gregg et al., 2021*).

An important driver for ECS at fertility clinics is the specific context of third-party reproduction, where for reproductive treatments such as IVF, donor spermatozoa or donor oocytes can be used (*Dondorp et al., 2014*). Following the current recommendations of the American Society for Assisted Reproduction, genetic testing of donors should, at least, include carrier testing for the same conditions as (following ACOG) those conditions screened for in those planning a pregnancy (*ASRM, 2021*). In Europe, where recommendations for universal carrier screening are generally lacking, the European Union Directive on human tissues and cells requires that gamete donors should be tested 'for autosomal recessive genes known to be prevalent [. . .] in the donor's ethnic background' (*Commission Directive, 2006*). Individual countries need to devise their own laws on how to reach goals specified in directives. To date, the Netherlands has not implemented carrier screening of gamete donors. While carrier testing of gamete donors remains limited to a few recessive disorders, following recommendations or legal requirements, candidate donors found to be carriers can be excluded from donating. However, this policy is no longer possible with ECS as, with ever-widening panels, all donors will inevitably turn out to be carriers of at least some autosomal-recessive disorders. Consequently, the step towards ECS of gamete donors can only be maintained if ECS-tested donors are matched with ECS-tested recipients with an eye to avoiding carrier couples. It is based on this strategy that some commercial fertility centres in Europe have started offering ECS in their practices, both to those depending on third party material and for couples using their own gametes (*Abulí et al., 2016*; *Gil-Arribas et al., 2016*; *Martin et al., 2015*).

There is still much debate about the conditions under which ECS could be implemented responsibly in a fertility context. With regards to medically assisted reproduction, concerns are that the drive towards more comprehensive testing may lead to ignoring the interests of donors, such as revealing genetic information that may be burdensome to themselves, or lead to societal discrimination and/or stigmatization, and put pressure upon recipients for what is, at best, a marginal added safety benefit (*Dondorp et al., 2014*; *Glenn et al., 2020*; *Mertes et al., 2018*). With regards to ECS for those intending to use their own gametes, the aim would be to provide them with options for reproductive decision-making, including preimplantation genetic testing (PGT) of embryos, prenatal diagnosis and the use of donor gametes in couples that turn out to be carrier couples. Here the debate is whether such carrier testing should be offered in this context, ahead of offering it to the general population of reproductive age. The Ethics Committee of the European Society of Human Reproduction and Embryology argued that there may be good reasons for doing so, especially in view of the fact that those coming for assisted reproduction are already proactively engaged with reproductive decision-making. However, in view of many remaining uncertainties, pilot projects would be needed to establish the conditions under which ECS could be offered responsibly in this context. These would have the added benefit of shedding light on the proportionality of offering ECS to all couples of reproductive age (*de Wert et al., 2021*).

Little is known about the attitudes of professionals in the field of medically assisted reproduction with regards to the pros and cons of offering ECS to their patients. This study aimed to explore the views of reproductive healthcare providers in the Netherlands, a country where assisted reproduction, including PGT for those at a known high risk of having a child with a serious genetic disorder, is reimbursed, where licensed clinics do not operate in a commercially competitive field, and where professional recommendations state that there should be no systematic carrier testing of gamete donors while such testing is not offered to the general population of reproductive age (*Netherlands Society of Obstetrics and Gynaecology and Dutch Association of Clinical Embryologists, 2018*).

MATERIALS AND METHODS

Study design

A qualitative study design using semi-structured interviews was used to capture a range of diverse arguments among reproductive healthcare providers regarding their potential involvement in offering ECS.

Setting and study population

In the Netherlands, DNA analysis legally requires a government license, and, at present, these licenses are only held by academic centres. This means there is no market for commercial providers of carrier tests, apart from foreign laboratories analysing material sent to them by Dutch clients. At present, carrier screening in the Netherlands is not common, although it is advised that it should be offered opportunistically to high-risk groups based on ancestry and/or consanguinity, based on a guideline developed by the Dutch Federation of Medical Specialists (2020). This includes screening based on ethnicity (ad-hoc haemoglobinopathy carrier screening and screening for disorders that are more common in the Ashkenazi Jewish community), geographic provenance (targeted testing within a Dutch genetically isolated population for founder mutations of five severe autosomal-recessive conditions) or consanguinity (whole-exome-based testing) (*Federation of Medical Specialists, 2020*; *Sallevelt et al., 2021*). Since 2016, ECS for 50–70 genes associated with severe autosomal-recessive disorders has been made available, but is rarely requested, for interested couples in the general population at their own cost at two university hospitals: Amsterdam University Medical Centre and University Medical Centre Groningen (*Delatycki et al., 2019*; *Plantinga et al., 2019*; *van Dijke et al., 2021*).

At the time of this study, there were an estimated 75 fertility clinics in the Netherlands, among which 15 were eligible to perform IVF (*Freya, 2020*). Licenses were held by eight academic medical centres, four peripheral facilities associated with academic centres, and three independent treatment centres (*Smeenk, 2019*). Due to the collectively funded healthcare system in the Netherlands, clinics are paid for their services by the Dutch Government, and do not have to compete with one another. Therefore, investigations and treatments offered are generally similar for all types of clinic, in contrast to countries with commercial or privatized healthcare settings.

Fertility clinics offer assisted reproduction to couples or individuals with fertility problems and/or at risk of having a child with a severe genetic condition. For both types of indication, this may involve the use of donor gametes. During a typical care trajectory, clients encounter numerous specialists, including obstetricians/gynaecologists, fertility doctors, nurse practitioners and clinical embryologists (*Choe et al., 2020*). For study inclusion, health professionals had to work directly with prospective parents at clinics with an IVF laboratory.

Sampling strategy

Initially, all 15 fertility clinics with an IVF laboratory were approached. Participants were identified through a purposive sampling strategy, and the first wave of recruitment consisted of six participants from three different clinics. Afterwards, snowball sampling took place for the recruitment of an additional 14 participants. During correspondence, the study was explained once more, and appointments were planned for the interviews. Interviews were held until data saturation was reached and no new themes emerged.

In total, 31 reproductive healthcare providers were contacted, of whom 20 (65%) participated in this study between May 2020 and January 2021. Twelve of the 15 clinics with licences in the Netherlands participated in the study. Several of those who chose not to participate stated that they did not want to participate due to time constraints or a disinterest in the study topic. Interviews were held online or by telephone with 11 gynaecologists, seven fertility doctors, one clinical embryologist and one nurse practitioner, and lasted for 45 min on average. There was no professional or personal connection between the interviewer (D.K.) and the healthcare providers interviewed. Work experience in the field of reproduction ranged from 1 to 37 years. The characteristics of the study population are listed in TABLE 1.

TABLE 1 CHARACTERISTICS OF THE STUDY POPULATION (N = 20)

		n (%)
Gender	Male	7 (35)
	Female	13 (65)
Clinic type	Academic	13 (65)
	Peripheral	4 (20)
	Independent	3 (15)
Specialism	Gynaecologist	11 (55)
	Fertility doctor	7 (35)
	Nurse practitioner	1 (5)
	Clinical embryologist	1 (5)
Years of experience	<10	6 (30)
	10–20	7 (35)
	>20	7 (35)

Data collection

An interview guide (see online supplementary material) was designed by the multidisciplinary research team, including health scientists, researchers in the field of carrier screening and a physician working in clinical genetics. Each topic covered in the interview included follow-up questions to probe more deeply into the interviewees' reasoning. One pilot interview was conducted at first to test the interview guide, and minor adjustments were made to clarify the questions. Interviews were recorded after approval and signing of the informed consent form. Audio recordings were deleted after transcription. Generic quotes were sent to the interviewees for member-checking. Due to the coronavirus 2019 (COVID-19) pandemic, it was impossible to perform live interviews; as such, interviews were conducted by telephone or online depending on the interviewees' preferences.

Data analysis

The interviews were transcribed verbatim. Transcripts were coded deductively using ATLAS.ti software for the generation and identification of themes. Subsequently, a coding guide was developed iteratively during data collection for further improvement of the overall reliability of data analysis. An initial set of transcripts (n = 4) was coded independently by two researchers (D.K. and I.D.) to increase internal validity through confirmed findings. Discrepancies were discussed until consensus was reached. The most relevant quotes were translated from Dutch to English, while trying to keep the reasoning of the quotes close to their context. Results were generated on the basis of arguments being mentioned frequently, conflicting arguments (i.e. deviant cases), and arguments considered to strongly affect outcome.

Ethical consideration

This study was reviewed by the Medical Ethics Committee of Amsterdam Medical Centre (Project No. W18_054 # 20.231, 7 May 2020), and deemed exempt from further review because the Act of Medical Research Involving Human Subjects (WMO) did not apply.

RESULTS

Results were clustered into four main themes with regards to reproductive healthcare providers' involvement in offering ECS at fertility clinics in the Netherlands: (i) general familiarity with and attitudes towards (involvement in) ECS; (ii) arguments in favour of involvement; (iii) arguments against involvement; and (iv) perceived necessities for future involvement. The findings are presented below.

General familiarity with and attitudes towards (involvement in) ECS

Familiarity with ECS

Participants differed in their experiences with carrier testing in general. Those who were familiar with carrier testing in their practice reported referring high-risk couples to clinical genetics to be counselled about the possibilities based on indication, for example due to a positive family history (e.g. thalassaemia, cystic fibrosis) or consanguinity. Although participants were aware of international (commercial) ECS offers, they were not involved in offering ECS, and knowledge was limited. Few participants were aware of the two ECS initiatives for interested couples in the general Dutch population:

> Yes, I have heard of it [ECS offers in the Netherlands], I think it was on the news 2 years ago. [R12 – Gynaecologist]

> Well, I find this [ECS offer at a nearby location] surprising, because we cooperate closely and the existence of that ECS offer is not known to me. [R17 – Fertility Doctor]

> To be honest, I had only read into it [ECS in the Netherlands] after you announced your research. [R10 – Gynaecologist]

One respondent was aware of ECS in the context of assisted reproduction. This respondent mentioned that their clinic offered gamete donors from foreign donor banks that went through ECS as a purchasable option:

> If they want to become pregnant via a donor, we usually offer the option to choose one from the [country name] donor bank [. . .]. With this, they can get a general screening panel which includes 300–400 of the most prevalent recessive conditions or get a specific donor screened to assess whether they are not a mismatch. [R13 – Fertility Doctor]

Attitudes towards offering ECS to certain groups

Participants discussed the possibility of offering ECS to certain groups, such as clients using donor gametes; all fertility clients ('general-risk couples'); or clients at increased risk, such as consanguineous couples. Participants indicated that offering ECS could be proportionate for general-risk couples interested in donor-assisted reproduction. Participants believed that by offering ECS to both gamete donors and recipients, matches between carriers of the same condition could be prevented:

> If you have not screened the donor and he or she turns out to have a certain genetic abnormality that the acceptor also has, then it is a risk we have all taken together. However, it would be nice if we could offer something to prevent that. [R10 – Gynaecologist]

Participants mentioned that if ECS was used to screen gamete donors, recipients would also need to be screened in order to allow matching aimed at avoiding carrier couples:

> If you are going to screen the donor, you will also have to screen the acceptor. [R10 – Gynaecologist]

> If a certain donor would not match one of our [screened] receivers, then we should not use it. In this case, we could just use a different donor from the bank. [R8 – Gynaecologist]

However, participants also believed that making ECS mandatory for gamete donors could result in donors being scared off from donating, for fear of being confronted with unwanted outcomes of broad-scope testing:

> The moment you start offering this kind of carrier test [ECS] in donor-assisted reproduction, you will have to start testing everything on a very large scale. [R1 – Gynaecologist]

> We already have a shortage of sperm donors. If you oblige potential donors to conduct genetic testing, which can lead to unexpected findings, it may result in deterrence of potential donors. [R12 – Gynaecologist]

Some participants believed an offer of ECS to be proportionate for all clients using IVF (general-risk couples), as so much work went into having a child:

> It feels terrible when so much medical work has already been put into having a child and it turns out to be unhealthy. Because those embryos [in assisted reproductive technology] are so valuable, it might be relevant to know whether that child is healthy for other conditions [in ECS]. But it already takes a lot of work to get a pregnancy done here. [R17 – Fertility Doctor]

However, most participants did not support offering ECS to couples who were having PGT. It was believed that offering ECS could heavily impact their chances of having a child by diminishing the number of potential embryos for transfer:

> There [with PGT], of course, you also have another problem. You will have the problem that the more you test, the more embryos fall off. You have to be very careful that you do not end up in a

situation where there are no more suitable embryos. [R9 – Gynaecologist]

Primarily consanguineous couples were believed to benefit from ECS, due to their increased risk of conceiving a child with an autosomal-recessive disorder:

> Well, I think we should initially focus on [screening] consanguineous couples, because there the risks are considerable and [societal] economic gains can also be made there. [R16 – Clinical Embryologist]

Whereas most participants did not believe that ECS had added value over the (targeted) panels that are already available for high-risk groups, some participants mentioned that offering ECS at their clinics could still be interesting. In these cases, offering ECS opportunistically by expanding targeted panels of high-risk couples seemed logical to them, considering that these groups already undergo specific carrier tests:

> We are already testing them on basis of a special indication. Simply expanding that [targeted panel] seems to me as the next logical step. [R9 – Gynaecologist]

It was also believed that these groups would likely experience negligible difficulties in understanding ECS during counselling, due to their familiarity with the subject matter:

> I can imagine that with high-risk groups, you are able to have an open conversation about this [ECS] sooner. Those people often know better what they are talking about. [R19 – Gynaecologist]

Attitudes towards their own involvement in ECS

Most participants had no intention to embed ECS as a standard offer to all clients at their clinic in the near future. Reasons for this are reflected in the following sections covering both arguments in favour and against involvement in offering ECS. Nevertheless, most participants were willing to participate in ECS-related implementation pilot studies:

> I would not yet offer it as part of routine care, but if you were to ask me if I would like to perform a pilot at our clinics [. . .], then I would definitely get involved. [R5 – Gynaecologist]

Meanwhile, participants predicted that ECS would become standard practice in the future, and that they might have to get involved in ECS at some point:

> I think this [ECS involvement] is a development we just cannot stop. [R4 – Gynaecologist]

> I only see it [ECS implementation] growing, so I think that it will gradually be implemented more and more in the future. [R11 – Gynaecologist]

Some participants feared that offering ECS as a standard practice could eventually create a 'slippery slope' towards a future of eugenics:

> I am afraid this might end up as a classic story of genetic selection; what is still okay? This raises some ethical concerns with me, and many others, I think. [R10 – Gynaecologist]

Arguments in favour

Opportunities offered by the setting

All participants agreed that fertility clinics could form a relevant setting for the provision of information regarding ECS, particularly as their patient population is already proactive towards reproductive decision-making before pregnancy:

> It is a place where you can already inform patients on all their options [regarding having children], this could include ECS. [R6 – Nurse Practitioner]

From a practical viewpoint, participants also believed that fertility clinics might form a suitable setting due to working directly with couples who might be interested in ECS:

> These people are already visiting the clinic, so you could just offer it as an addition. They are already right in front of you, so why not implement it anyway? [R5 – Gynaecologist]

Motivation to assist in reproduction and prevent suffering

Participants mentioned two forms of responsibility in their role as reproductive healthcare providers. On the one hand, their professional task was seen by many as assisting in reproduction. On the other hand, some participants stated that avoiding suffering due to severe autosomal-recessive conditions in infants born after their professional involvement by offering ECS could be proportional:

> It [ECS] obviously has pros and cons, but I do believe that the pros outweigh the cons in view of our [reproductive healthcare providers] pursuit in bringing healthy children into the world. [R9 – Gynaecologist]

Fertility clinics as a regulated setting to counter commercialization offers

Participants argued that ECS is already being offered (commercially) in countries outside the Netherlands, and that offering ECS at their clinics could counter commercialization for Dutch couples:

> Abroad, such [commercial] clinics will prosper from the possibility of offering tests which screen for extra conditions. They will just make it [couples' trajectory] cost an additional €5000. [R5 – Gynaecologist]

> What it is really all about, is whether we want to offer ECS under controlled settings [fertility clinics] to counter commercialization. [R1 – Gynaecologist]

It was generally feared that commercial offers of ECS would not include proper counselling. One participant reported direct experience, on several occasions, of patients fearfully asking how to interpret the results of their online-ordered ECS tests due to insufficient counselling:

> People just go to some clinic or even order these tests online. And afterwards they come to me and ask: what now doctor? [R5 – Gynaecologist]

Another argument against commercialization was that ECS in an unregulated setting was unfair for couples, as it could capitalize on the fear of conceiving an affected child:

> You are abusing the fears of people, and people are willing to put a lot of money for that. I think that is unfair. [R4 – Gynaecologist]

Arguments against

Lack of knowledge and familiarity with offering ECS

Overall, participants believed that their capabilities were currently insufficient for proper involvement in offering ECS. It was stated that there should be more information and education on how to

counsel, guide patients and make relevant choices if ECS were to be implemented:

> We would have to be more informed on this subject. How should we counsel and how do we provide proper guidance? And how would we make a choice afterwards? [R3 — Fertility Doctor]

Participants believed that, without support and education from clinical genetics, they would be incapable of offering ECS at their clinics. It was considered that clinical geneticists have the necessary knowledge for this, and could also help to support reproductive healthcare providers with, preferably uniform nationwide, materials:

> I think that the genetics department should develop an excellent leaflet. [R6 — Nurse Practitioner]

> I would first like to embed this with my collaboration with the clinical geneticists. [R8 — Gynaecologist]

> How will counselling and guidance be given? Because it has to be uniform nationally. [R2 — Fertility Doctor]

Insufficient staff and resources and potential high costs for clinics and/or couples

Participants felt that clinics had insufficient staff to embed ECS at their clinics at present:

> We are already too busy at the department, even with our usual procedures [. . .]. If we implement this, we would not be able to handle it all at the same time. [R1 — Gynaecologist]

This was linked to the financial state of fertility clinics, implying that there were insufficient resources to offer ECS at their clinics. Due to the expected high costs of ECS, participants mentioned that offering ECS would likely impact their budget significantly. One respondent even feared that ECS involvement would lead to sacrificing other services at the clinic:

> I imagine this could lead to sacrificing something else, because we will not get more money for this extra work. [R2 — Fertility Doctor]

Participants mentioned that in order to avoid an impact on the budget of fertility clinics, couples would have to pay out-of-pocket for the service. Many low-income couples would not be able to meet these costs, which would probably lead to more unequal access:

> It would mean that you are excluding a part of the population, so that only the rich and famous can get tested. [R2 — Fertility Doctor]

It was mentioned that the Dutch National Healthcare Institute should take responsibility for funding ECS in order to create a situation where society could eventually capitalize on embedding ECS at clinics:

> I think ZIN [Zorg Instituut Nederland; Dutch National Healthcare Institute] could take the responsibility for initially funding this [ECS]. Because, we could possibly economize on preventing the life-time costs of all these diseases. [R5 — Gynaecologist]

Emotional impact that ECS may have on couples

Participants believed that their involvement in offering ECS could have adverse consequences for their patients. It was suggested that having ECS could give a false sense of reassurance:

> People keep thinking they can get guarantees regarding the health of their children, there are no guarantees for this. [R6 — Nurse Practitioner]

Meanwhile, it was also believed that some patients could experience a disproportionate amount of fear by receiving positive (unfavourable) test results in irresponsibly composed panels:

> I think it [ECS test results] might bring things up which cannot be treated or do not even need to be treated, which might still worry people enormously. [R3 — Fertility Doctor]

Participants also feared that offering ECS at their clinic might impact identified carriers financially. The possibility of health insurers taking ECS results into account when defining medical insurance plans was (falsely) seen as a possible threat:

> I think it might even have consequences for the rest of your life, due to the involvement of insurers or for mortgages and things like that. [R7 — Gynaecologist]

Perceived complexity of counselling and expected elongation of waiting lists

Participants believed that couples might have trouble interpreting pre-test education and ECS results without pre- and post-test counselling. It was expected that this would likely lead to a need for longer, more frequent and complex counselling:

> Counselling would be really complex, because it is very complicated stuff. It would be hard to dose properly for patients. [R14 — Fertility Doctor]

Meanwhile, participants were in favour of offering ECS accompanied by thorough counselling, as opposed to fast tests with incomplete counselling:

> I think ECS has to be given with really good counselling. But I do not think this will be easy, and I really see a danger in the idea of offering quick tests with limited counselling. [R14 — Fertility Doctor]

Participants also anticipated that offering ECS to couples at their clinic would take up a substantial amount of their time, which would lead to unacceptably long waiting lists:

> At the moment, we believe we should offer this [ECS] to a broader part of the population [couples at fertility clinics], it would lead to unacceptably long waiting lists. [R1 — Gynaecologist]

Expected low impact of ECS on reducing the burden of diseases

Participants doubted the impact of ECS on a larger scale due to the perceived low prevalence of recessive conditions:

> I do not think this [ECS] will change our world [. . .]. The chance of conceiving an affected child is extremely low. [R1 — Gynaecologist]

Meanwhile, participants believed that ECS would only reduce the risk of conceiving a child with (congenital) diseases to a small extent:

> When couples decide to have a child, they have an estimated 3% chance of conceiving a child with a congenital disease [. . .], that is the normal background risk. The diseases we are talking about [when screening with ECS] are super rare, so you are reducing this

> 3% by a minimum amount. [R4 – Gynaecologist]

According to participants, implementing ECS at fertility clinics would have a relatively small impact on the knowledge and use of ECS in the whole population. The primary argument was that fertility clinics only see a marginal proportion of the (relevant) population who wish to have a child:

> People that visit our clinic are only 10% of the couples with a child wish. The other 90% get pregnant within a year, so they do not visit us. [R2 – Fertility Doctor]

Another factor that made the participants doubt the impact of offering ECS at their clinics was that they had experienced little demand for ECS among their patients, with very few requests for it:

> From my experience, the demand [for ECS] is pretty low, which is why I also think the impact would not be that big. [R15 – Gynaecologist]

Other reasons reported were that the patients' main request is getting pregnant or a lack of awareness of ECS. However, participants also believed that the population is getting more informed as well as interested in ECS, which could facilitate demand at some point. The uncertainty revolving around the implementation of ECS was evident, but the importance of debate was never questioned:

> There is a lot of public debate going on around it [ECS] now, I think that is facilitating it [implementation] enormously. The momentum seems quite good, and I do see a future in it, but we should be careful using words like 'never', 'always', 'everyone' or 'now'. [R16 – Clinical Embryologist]

Necessities for future involvement

Participants stated that they would need more evidence to substantiate their potential involvement in offering ECS, such as by getting involved in ECS through pilot implementation studies. It was noted that involvement would only take place after studying three important factors. Firstly, as participants were unaware of the costs of ECS and what could be saved in healthcare expenditures, the provision of a cost–benefit overview was deemed to be crucial:

> I really think a cost–benefit overview should be made. One that shows how much the strategy [of implementing ECS] costs in contrast to what we currently spend on treating [children with] genetic diseases. [R2 – Fertility Doctor]

Secondly, participants addressed the importance and lack of studies covering the implications, for clinics and patients, of implementing ECS at fertility clinics:

> In order to consider involvement with this [ECS], you really have to prove the collective value, and if it outweighs practice and society related risks. But we are not that far yet. We obviously do not know all the consequences this might have. [R20 – Gynaecologist]

Finally, it was mentioned that the overall demand and support for ECS from the patients' perspective should be assessed thoroughly prior to implementation:

> I think the most important thing is to see if our patients would even want it. [. . .] If a large group of patients does not want it, then we probably should not invest our time and energy into it. [R7 – Gynaecologist]

DISCUSSION

While agreeing that the field of medically assisted reproduction provides a unique opportunity for offering ECS, reproductive healthcare workers feel a lack of capability and limited motivation to offer ECS at their clinics. They reported insufficient awareness and familiarity with ECS, and therefore found it difficult to estimate its added value for couples visiting fertility clinics. This, together with the perceived complexity of counselling for ECS, was the reason why participants felt a lack of capability to offer it at their clinics. They also highlighted obstacles related to the funding and acquisition of staff for ECS involvement. These difficulties notwithstanding, they thought that, in principle, they were well placed to inform or offer ECS opportunistically to high-risk couples visiting their clinics, partly because they expected them to be familiar with the subject matter. They had the highest motivation to offer ECS to consanguineous couples, gamete donors and those in need of donor gametes. Meanwhile, participants doubted whether the pros could outweigh the cons of offering ECS at their fertility clinics to general-risk couples with an indication for assisted reproductive technology alone. Participants generally requested more evidence of the added value of offering ECS in assisted reproductive technology in terms of cost–benefit, implications for both the clinic and patient populations, and whether their patient population would be interested in ECS. Participants seemed willing to participate in ECS-related pilot studies for the generation of this evidence.

The arguments related to capability, opportunity and motivation that professionals mentioned are three key elements that influence behaviour, forming the basis of theoretical behavioural change models (*Atkins et al., 2017*). These models have been used in a previous interview study among health professionals (primarily general practitioners) in Australia, showing that lack of knowledge is an important barrier for engaging with screening (*Best et al., 2023*). The lack of capability, such as limited familiarity and insufficient experience with carrier screening among reproductive healthcare providers in Europe, has been described in previous studies (*Dungan, 2018*; *Holtkamp et al., 2017*; *Janssens et al., 2017*). Similar to the present findings, *Cho et al. (2013)* stated that without proper education, reproductive healthcare providers might experience difficulties in offering ECS at fertility clinics, especially with regards to communicating ECS results to patients. In a Swedish study, healthcare professionals, including clinical geneticists, expected reproductive healthcare providers to experience issues with counselling and the provision of dosed information to couples (*Matar et al., 2016*). Participants, including clinical geneticists, agreed that reproductive healthcare providers would need more education on the provision of ECS. *Schuurmans et al. (2019)* demonstrated that offering ECS through educated general practitioners was feasible for a population-based screening approach.

Participants were not in favour of introducing out-of-pocket costs to couples to make ECS involvement possible, as it would lead to unequal access to ECS. The findings also suggest that fertility clinics would need adjusted budget allocation structures or, preferably, funding to avoid inequity among couples. Previous studies have shown how socio-economic status and costs may play a role in the uptake of ECS (*Bajaj and Gross, 2014*; *Robson et al.,*

2020; Van Steijvoort et al., 2020). The potential saving of healthcare costs of children with genetic diseases was raised by the participants, and the need for evaluation of costs and benefits, although costs may affect their own services while savings might be achieved elsewhere.

With regards to demand, numerous studies have assessed the general public's interest in ECS. Ong et al. (2018) showed that more than two-thirds of Australian couples would be interested in ECS. Meanwhile, in both Sweden and the Netherlands, approximately one-third of (parent) couples from the general population would consider an offer of ECS (Ekstrand Ragnar et al., 2016; Nijmeijer et al., 2019). Few studies have covered interest in ECS specifically among fertility clinic visitors, although ECS is offered increasingly in this setting (Capalbo et al., 2021). One US study showed that uptake of ECS was higher among women who had been counselled in preparation for IVF compared with women who had been counselled for other reasons (Larsen et al., 2019). Another US study showed that in a survey of 100 participants considering or receiving fertility treatment, 34% were positive about ECS (Pereira et al., 2019).

Arguments related to motivation were more diverse among reproductive healthcare providers.

Discrepancies which may have affected the participants' motivation to offer ECS in reproductive health care include the perceived prevalence of autosomal-recessive conditions, and the expected clinical or added value of ECS in reducing the chance of having a child with a congenital or genetic disorder in the first place. For instance, the chance of conceiving a child affected by an autosomal-recessive condition is closely comparable with that of aneuploidy. In the Netherlands, all pregnant women are offered non-invasive prenatal testing (NIPT), while the general average risk of conceiving a child with Down syndrome in Europe is approximately one in 400 to one in 600 (de Graaf et al., 2020; de Groot-van der Mooren et al., 2021). The argument that the probability of conceiving an affected child is too low has not been decisive for offering NIPT in the Netherlands. Studies in other countries have shown that ECS can have significant value for patients seeking assisted reproductive technology, and that genes appropriate for screening can be identified for different populations (Capalbo et al., 2021; Chen et al., 2023; Xi et al., 2020).

Participants saw little justification for offering ECS in reproductive health care alone, as one would only reach a minority of those for whom such an offer might be useful. This also links in with concerns about justice, as discussed in the literature (de Wert et al., 2021). On the other hand, as this is one of the few healthcare settings attended by couples before pregnancy, pilots in this context may also be considered as a means to generate information relevant to decision-making about ECS as an offer to the general population.

Limitations of the study

This study provided rich insights into the perspectives of a diverse group of reproductive healthcare providers working at fertility clinics in the Netherlands. The study had several limitations. First, snowball sampling was used for recruitment; this poses a possible threat to data variation, as new participants may share similar views towards the subject. Similarly, numerous invitees did not wish to participate due to disinterest in the study. Furthermore, most participants worked at academic medical centres, as opposed to peripheral facilities or independent treatment centres. Finally, due to the COVID-19 pandemic, some interviews were conducted by telephone, potentially limiting non-verbal communication between the researcher and participant.

Implications for further research and practice

In the Netherlands, besides carrier screening for high-risk groups based on ancestry, geographic origin or consanguinity, ECS for the general population is only available at their own cost in some local initiatives, and general awareness remains low (Holtkamp et al., 2017). Elsewhere, settings where ECS could be offered are still being explored and pilot tested, such as in Australia (Delatycki et al., 2019; Kirk et al., 2021; Rowe and Wright, 2020). In general, ECS shows an increasingly heterogeneous landscape of screening offers, driven by commercial companies (Chokoshvili et al., 2017; Kraft et al., 2019).

This study provides insights into the capability, opportunity and motivation of reproductive healthcare professionals regarding offering ECS in the Netherlands, making it relevant for both ECS development and fertility clinics throughout the Netherlands. To address this potential, four recommendations have been formulated.

First, reproductive healthcare providers felt a lack of capability towards offering ECS or counselling couples, whether before or after the actual offer. It was advocated that there would be a need for proper education and training prior to involvement. Guidelines would also be needed if ECS was to be offered systematically throughout the country at some point.

Second, some respondents mentioned that their clinics offered the opportunity to provide ECS to certain groups. Consanguineous couples were seen as the main target group, followed by other high-risk couples. Opinions were relatively positive regarding an offer of ECS to general-risk couples interested in donor gametes. Opinions differed on how this should be implemented, especially on the choice between initially testing gamete donors versus gamete recipients. An opportunistic offer of ECS to general-risk couples with an indication for assisted reproductive technology was not seen as proportionate. This shows that ongoing debate regarding ECS in assisted reproduction, whether through donors or IVF, is needed for policy development, and should include the definition of target groups. To avoid complex counselling about individual carrier test results in ECS, a 'couple-based' approach could be considered where the donor and recipient are tested in parallel, and a positive result is only reported when the donor and recipient are carriers of the same (severe) autosomal-recessive disorder, because, in that case, the risk of conceiving a child affected with that disorder increases by 25% (Plantinga et al., 2019; Schuurmans et al., 2019).

Third, participants were not entirely motivated regarding involvement with ECS before having an overview of the costs and benefits, as well as a system for reimbursement for offering ECS at fertility clinics. To answer participants' questions on costs and benefits, all variables should be taken into consideration, including the involvement of all relevant stakeholders, staff, training, counselling, creating awareness, the tests themselves, and potential savings in healthcare costs related to the recessive conditions included in these tests.

The final recommendation would be for fertility clinics to participate in further pilot studies on ECS. As discussed previously, studies on ECS implementation are still scarce, suggesting that there is still a lot to learn about offering ECS in this context, and also with regards to decision-making about offering ECS to the general population. By initiating pilot studies, clinics will also be able to evaluate the interest among visiting couples, including the perspectives of gamete donors; their eventual involvement in offering ECS; and provide relevant insights for other clinics, including in other countries and settings.

DATA AVAILABILITY

The data that has been used is confidential.

ACKNOWLEDGEMENTS

The authors wish to thank all the involved reproductive healthcare providers and fertility clinics for their contribution and time for this study.

FUNDING

L.H., I.v.L., W.D., P.L. and M.C. received funding from the Netherlands Organization for Health Research and Development (ZonMw) to study the societal and ethical aspects of preconception carrier screening (Grant No. 854020002). The salary of I.D. was funded by the Amsterdam Reproduction and Development Research Institute (Grant No. 2017–170116).

AUTHOR CONTRIBUTIONS

M.C., L.H., D.K. and I.D. initiated the study and developed the interview guide. D.K., M. C., W.D. and P.L. were involved in recruitment. D.K. conducted the interviews and coded all transcripts. A subset of transcripts ($n = 4$) was also coded by I.D. D. K. drafted the initial manuscript. All authors reviewed the manuscript, critically revised it for important intellectual content, and approved the final version.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2024.103857.

REFERENCES

Abulí, A., Boada, M., Rodríguez-Santiago, B., Coroleu, B., Veiga, A., Armengol, L., Barri, P.N., Pérez-Jurado, L.A., Estivill, X., 2016. NGS-Based assay for the identification of individuals carrying recessive genetic mutations in reproductive medicine. Hum Mutat 37, 516–523. https://doi.org/10.1002/humu.22989.

ACOG, American College of Obstetrics and Gynecology, 2017. Committee Opinion No. 691: Carrier screening for genetic conditions. Obstetrics & Gynecology 129.

ASRM, American Society for Reproductive Medicine, 2021. Guidance regarding gamete and embryo donation. Fertil Steril 115, 1395–1410. https://doi.org/10.1016/j.fertnstert.2021.01.045.

Atkins, L., Francis, J., Islam, R., O'Connor, D., Patey, A., Ivers, N., Foy, R., Duncan, E.M., Colquhoun, H., Grimshaw, J.M., Lawton, R., Michie, S., 2017. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implementation Science 12, 77. https://doi.org/10.1186/s13012-017-0605-9.

Bajaj, K., Gross, S.J., 2014. Carrier screening: past, present, and future. J Clin Med 3, 1033–1042. https://doi.org/10.3390/jcm3031033.

Best, S., Long, J.C., Fehlberg, Z., Theodorou, T., Hatem, S., Archibald, A., Braithwaite, J., 2023. The more you do it, the easier it gets: using behaviour change theory to support health care professionals offering reproductive genetic carrier screening. European Journal of Human Genetics 31, 430–444. https://doi.org/10.1038/s41431-022-01224-5.

Capalbo, A., Fabiani, M., Caroselli, S., Poli, M., Girardi, L., Patassini, C., Favero, F., Cimadomo, D., Vaiarelli, A., Simon, C., Rienzi, L.F., Ubaldi, F.M., 2021. Clinical validity and utility of preconception expanded carrier screening for the management of reproductive genetic risk in IVF and general population. Human Reproduction 36, 2050–2061. https://doi.org/10.1093/humrep/deab087.

Chen, H.-Y., Lin, S.-Y., Shih, J.-C., Kang, J., Tai, Y.-Y., Shaw, S.W., Chen, K.-C., Mai, K., Lee, C.-N., 2023. Changing the standardised obstetric care by expanded carrier screening and counselling: a multicentre prospective cohort study. J Med Genet. https://doi.org/10.1136/jmg-2023-109268.

Cho, D., McGowan, M.L., Metcalfe, J., Sharp, R.R., 2013. Expanded carrier screening in reproductive healthcare: Perspectives from genetics professionals. Human Reproduction 28, 1725–1730. https://doi.org/10.1093/humrep/det091.

Choe, J., Archer, J.S., Shanks, A.L., 2020. Vitro Fertilization. StatPearls Publishing, Treasure Island (FL), Indiana University. PMID: 32965937.

Chokoshvili, D., Vears, D.F., Borry, P., 2017. Growing complexity of (expanded) carrier screening: Direct-to-consumer, physician-mediated, and clinic-based offers. Best Pract Res Clin Obstet Gynaecol 44, 57–67. https://doi.org/10.1016/j.bpobgyn.2017.02.006.

Commission Directive, 2006. Commission Directive 2006/17/EC on implementing Directive 2004/23/EC of the European Parliament and of the Council as regards certain technical requirements for the donation, procurement and testing of human tissues and cells.

de Graaf, G., Buckley, F., Skotko, B., 2020. Estimation of the number of people with Down syndrome in Europe. European Journal of Human Genetics 29, 1–9. https://doi.org/10.1038/s41431-020-00748-y.

de Groot-van der Mooren, M., de Graaf, G., Weijerman, M.E., Hoffer, M.J.V, Knijnenburg, J., van der Kevie-Kersemaekers, A.-M.M.F., Kooper, A.J.A., Voorhoeve, E., Sikkema-Raddatz, B., van Zutven, L.J.C.M., Srebniak, M.I., Huijsdens-van Amsterdam, K., Engelen, J.J.M., Smeets, D., van Kaam, A.H., Cornel, M.C., 2021. Does non-invasive prenatal testing affect the livebirth prevalence of Down syndrome in the Netherlands? A population-based register study. Prenat Diagn 41, 1351–1359. https://doi.org/10.1002/pd.6003.

de Wert, G., van der Hout, S., Goddijn, M., Vassena, R., Frith, L., Vermeulen, N., Eichenlaub-Ritter, U., 2021. ESHRE Ethics Committee, 2021. The ethics of preconception expanded carrier screening in applicants of medically assisted reproduction. Hum Reprod Open hoaa063. https://doi.org/10.1093/hropen/hoaa063.

Delatycki, M., Alkuraya, F., Archibald, A., Castellani, C., Cornel, M., Grody, W., Henneman, L., Ioannides, A., Kirk, E., Laing, N., Lucassen, A., Massie, J., Schuurmans, J., Thong, M.-K., Langen, I., Zlotogora, J., 2019. International perspectives on the implementation of reproductive carrier screening. Prenat Diagn 40, 301–310. https://doi.org/10.1002/pd.5611.

Dondorp, W., De Wert, G., Pennings, G., Shenfield, F., Devroey, P., Tarlatzis, B., Barri, P., Diedrich, K., Eichenlaub-Ritter, U., Tüttelmann, F., Provoost, V., 2014. ESHRE Task Force on Ethics and Law 21: genetic screening of gamete donors: ethical issues†. Human Reproduction 29, 1353–1359. https://doi.org/10.1093/humrep/deu111.

Dungan, J., 2018. Expanded carrier screening: what the reproductive endocrinologist needs to know. Fertil Steril 109, 183–189. https://doi.org/10.1016/j.fertnstert.2017.11.030.

Ekstrand Ragnar, M., Tydén, T., Kihlbom, U., Larsson, M., 2016. Swedish parents' interest in preconception genetic carrier screening. Ups J Med Sci 121, 289–294. https://doi.org/10.1080/03009734.2016.1218575.

Federation of Medical Specialists, 2020. Richtlijn Preconceptie dragerschapsonderzoek (PDO) voor hoogrisicogroepen. [Dutch Guideline, WWW Document] URL https://richtlijnendatabase.nl/richtlijn/preconceptie_dragerschapsonderzoek_pdo_voor_hoogrisicogroepen/startpagina_-_pdo.html (accessed 1.12.20)

Freya, 2020. Kies de kliniek die bij je past [WWW Document]. URL https://www.freya.nl/kinderwens/starters-in-het-medisch-circuit/kliniek-keuze/(accessed 4.12.20).

Fridman, H., Yntema, H.G., Mägi, R., Andreson, R., Metspalu, A., Mezzavilla, M., Tyler-Smith, C., Xue, Y., Carmi, S., Levy-Lahad, E., Gilissen, C., Brunner, H.G., 2021. The landscape of autosomal-recessive pathogenic variants in European populations reveals phenotype-specific effects. American Journal of Human Genetics 108, 608–619. https://doi.org/10.1016/j.ajhg.2021.03.004.

Gil-Arribas, E., Herrer, R., Serna, J., 2016. Pros and cons of implementing a carrier genetic test in an infertility practice. Curr Opin Obstet Gynecol 28, 172–178. https://doi.org/10.1097/GCO.0000000000000272.

Glenn, T.L., Pereira, N., Madeira, J., Maxwell, R.A., Parry, J.P., Mertes, H., Pennings, G., Lindheim, S.R., 2020. Reproductive

endocrinology infertility (REI) specialists' utilization and attitudes toward expanded carrier screening (ECS) for third-party oocyte donors. Journal of Obstetrics and Gynecology of India 70, 409–411. https://doi.org/10.1007/s13224-019-01297-1.

Goossens, J., De Roose, M., Van Hecke, A., Goemaes, R., Verhaeghe, S., Beeckman, D., 2018. Barriers and facilitators to the provision of preconception care by healthcare providers: A systematic review. Int J Nurs Stud 87, 113–130. https://doi.org/10.1016/j.ijnurstu.2018.06.009.

Gregg, A.R., Aarabi, M., Klugman, S., Leach, N.T., Bashford, M.T., Goldwaser, T., Chen, E., Sparks, T.N., Reddi, H.V, Rajkovic, A., Dungan, J.S., 2021. Screening for autosomal recessive and X-linked conditions during pregnancy and preconception: a practice resource of the American College of Medical Genetics and Genomics (ACMG). Genetics in Medicine 23, 1793–1806. https://doi.org/10.1038/s41436-021-01203-z.

Gregg, A.R., Edwards, J.G., 2018. Prenatal genetic carrier screening in the genomic age. Semin Perinatol 42, 303–306. https://doi.org/10.1053/j.semperi.2018.07.019.

Henneman, L., Borry, P., Chokoshvili, D., Cornel, M.C., Van El, C.G., Forzano, F., Hall, A., Howard, H.C., Janssens, S., Kayserili, H., Lakeman, P., Lucassen, A., Metcalfe, S.A., Vidmar, L., De Wert, G., Dondorp, W.J., Peterlin, B., 2016. Responsible implementation of expanded carrier screening. European Journal of Human Genetics. https://doi.org/10.1038/ejhg.2015.271.

Holtkamp, K.C.A., Vos, E.M., Rigter, T., Lakeman, P., Henneman, L., Cornel, M.C., 2017. Stakeholder perspectives on the implementation of genetic carrier screening in a changing landscape. BMC Health Serv Res 17, 146. https://doi.org/10.1186/s12913-017-2083-9.

Janssens, S., Chokoshvili, D., Vears, D.F., De Paepe, A., Borry, P., 2017. Pre- and post-testing counseling considerations for the provision of expanded carrier screening: exploration of European geneticists' views. BMC Med Ethics 18, 46. https://doi.org/10.1186/s12910-017-0206-9.

Kirk, E.P., Ong, R., Boggs, K., Hardy, T., Righetti, S., Kamien, B., Roscioli, T., Amor, D.J., Bakshi, M., Chung, C.W.T., Colley, A., Jamieson, R.V, Liebelt, J., Ma, A., Pachter, N., Rajagopalan, S., Ravine, A., Wilson, M., Caruana, J., Casella, R., Davis, M., Edwards, S., Archibald, A., McGaughran, J., Newson, A.J., Laing, N.G., Delatycki, M.B., 2021. Gene selection for the Australian Reproductive Genetic Carrier Screening Project ('Mackenzie's Mission'). Eur J Hum Genet 29, 79–87. https://doi.org/10.1038/s41431-020-0685-x.

Kraft, S.A., Duenas, D., Wilfond, B.S., Goddard, K.A.B., 2019. The evolving landscape of expanded carrier screening: challenges and opportunities. Genetics in Medicine. https://doi.org/10.1038/s41436-018-0273-4.

Larsen, D., Ma, J., Strassberg, M., Ramakrishnan, R., Van den Veyver, I.B., 2019. The uptake of pan-ethnic expanded carrier screening is higher when offered during preconception or early prenatal genetic counseling. Prenat Diagn 39, 319–323. https://doi.org/10.1002/pd.5434.

Martin, J., Asan, Yi, Y., Alberola, T., Rodríguez-Iglesias, B., Jiménez-Almazán, J., Li, Q., Du, H., Alama, P., Ruiz, A., Bosch, E., Garrido, N., Simon, C., 2015. Comprehensive carrier genetic test using next-generation deoxyribonucleic acid sequencing in infertile couples wishing to conceive through assisted reproductive technology. Fertil Steril 104, 1286–1293. https://doi.org/10.1016/j.fertnstert.2015.07.1166.

Matar, A., Kihlbom, U., Höglund, A.T., 2016. Swedish healthcare providers' perceptions of preconception expanded carrier screening (ECS)-a qualitative study. J Community Genet 7, 203–214. https://doi.org/10.1007/s12687-016-0268-2.

Mertes, H., Lindheim, S.R., Pennings, G., 2018. Ethical quandaries around expanded carrier screening in third-party reproduction. Fertil Steril 109, 190–194. https://doi.org/10.1016/j.fertnstert.2017.11.032.

Netherlands Society of Obstetrics and Gynaecology and Dutch Association of Clinical Embryologists, 2018. Landelijk standpunt spermadonatie Specifieke eisen voor spermadonoren. Utrecht. [WWW Document] URL https://www.nvog.nl/wp-content/uploads/2018/04/landelijk-standpunt-spermadonatie-KLEM-en-NVOG-april-2018.pdf (accessed 1.12.20)

Nijmeijer, S.C.M., Conijn, T., Lakeman, P., Henneman, L., Wijburg, F.A., Haverman, L., 2019. Attitudes of the general population towards preconception expanded carrier screening for autosomal recessive disorders including inborn errors of metabolism. Mol Genet Metab 126, 14–22. https://doi.org/10.1016/j.ymgme.2018.12.004.

Ong, R., Howting, D., Rea, A., Christian, H., Charman, P., Molster, C., Ravenscroft, G., Laing, N., 2018. Measuring the impact of genetic knowledge on intentions and attitudes of the community towards expanded preconception carrier screening. J Med Genet 55, 744–752. https://doi.org/10.1136/jmedgenet-2018-105362.

Pereira, N., Wood, M., Luong, E., Briggs, A., Galloway, M., Maxwell, R.A., Lindheim, S.R., 2019. Expanded genetic carrier screening in clinical practice: a current survey of patient impressions and attitudes. J Assist Reprod Genet 36, 709–716. https://doi.org/10.1007/s10815-019-01414-z.

Plantinga, M., Birnie, E., Schuurmans, J., Buitenhuis, A.H., Boersma, E., Lucassen, A.M., Verkerk, M.A., van Langen, I.M., Ranchor, A.V, 2019. Expanded carrier screening for autosomal recessive conditions in health care: Arguments for a couple-based approach and examination of couples' views. Prenat Diagn 39, 369–378. https://doi.org/10.1002/pd.5437.

Robson, S., Caramins, M., Saad, M., Suthers, G., 2020. Socioeconomic status and uptake of reproductive carrier screening in Australia. Australian and New Zealand Journal of Obstetrics and Gynaecology 60, 976–979. https://doi.org/10.1111/ajo.13206.

Rowe, C.A., Wright, C.F., 2020. Expanded universal carrier screening and its implementation within a publicly funded healthcare service. J Community Genet 11, 21–38. https://doi.org/10.1007/s12687-019-00443-6.

Sallevelt, S.C.E.H., Stegmann, A.P.A., de Koning, B., Velter, C., Steyls, A., van Esch, M., Lakeman, P., Yntema, H., Esteki, M.Z., de Die-Smulders, C.E.M., Gilissen, C., van den Wijngaard, A., Brunner, H.G., Paulussen, A.D.C., 2021. Diagnostic exome-based preconception carrier testing in consanguineous couples: results from the first 100 couples in clinical practice. Genetics in Medicine 23, 1125–1136. https://doi.org/10.1038/s41436-021-01116-x.

Schuurmans, J., Birnie, E., van den Heuvel, L.M., Plantinga, M., Lucassen, A., van der Kolk, D.M., Abbott, K.M., Ranchor, A.V, Diemers, A.D., van Langen, I.M., 2019. Feasibility of couple-based expanded carrier screening offered by general practitioners. European Journal of Human Genetics 27, 691–700. https://doi.org/10.1038/s41431-019-0351-3.

Smeenk, J.M.J., 2019. IVF-cijfers per centrum 2018.

van der Hout, S., Dondorp, W., de Wert, G., 2019. The aims of expanded universal carrier screening: Autonomy, prevention, and responsible parenthood. Bioethics 33, 568–576. https://doi.org/10.1111/bioe.12555.

van Dijke, I., Lakeman, P., Sabiri, N., Rusticus, H., Ottenheim, C.P.E., Mathijssen, I.B., Cornel, M.C., Henneman, L., 2021. Couples' experiences with expanded carrier screening: evaluation of a university hospital screening offer. European Journal of Human Genetics 29, 1252–1258. https://doi.org/10.1038/s41431-021-00923-9.

Van Steijvoort, E., Chokoshvili, D., W Cannon, J., Peeters, H., Peeraer, K., Matthijs, G., Borry, P., 2020. Interest in expanded carrier screening among individuals and couples in the general population: Systematic review of the literature. Hum Reprod Update 26, 335–355. https://doi.org/10.1093/humupd/dmaa001.

Xi, Y., Chen, G., Lei, C., Wu, J., Zhang, S., Xiao, M., Zhang, W., Zhang, Y., Sun, X., 2020. Expanded carrier screening in Chinese patients seeking the help of assisted reproductive technology. Mol Genet Genomic Med 8, e1340. https://doi.org/10.1002/mgg3.1340.

Received 30 September 2023; received in revised form 1 January 2024; accepted 29 January 2024.

RBMO

ARTICLE

Direct-to-consumer DNA testing: the perspectives and experiences of donor conceived young adults in the UK

BIOGRAPHY

Dr Sophie Zadeh is Reader in Family Psychology at the School of Psychology, University of Sussex. Her current research investigates the experiences of donor conceived people in the UK.

Sophie Zadeh[1,*,**]

KEY MESSAGE

Meanings ascribed to, and uses of, direct-to-consumer (DTC) DNA testing vary significantly among donor conceived young adults. There is both active interest and disinterest in DTC DNA testing among this group. The absence of support for those using DTC DNA tests should be addressed by practitioners, regulatory bodies and policymakers.

ABSTRACT

Research question: What meanings do donor conceived young adults give to direct-to-consumer DNA testing, and how does direct-to-consumer DNA testing relate to their lived experiences?

Design: Thirty-three young adults participated in in-depth interviews in November 2020 and September 2021 as part of a study of donor conceived people in the UK that focuses on the period of young adulthood. All participants were aged between 18 and 31 years, had been conceived by sperm donation at a time of legal donor anonymity, and were mainly resident in the UK. Interviews were analysed using reflexive thematic analysis.

Results: Nineteen participants (58%) had used at least one direct-to-consumer DNA test, and 14 (46%) had not. Three participants (9%) had learned about their donor conception inadvertently through a direct-to-consumer DNA test. Twelve participants (36%) had matched with their donor, someone conceived using the same donor, or both. Four related themes that capture participants' perspectives and experiences of direct-to-consumer DNA testing were identified: ruptures, disclosures, webs and temporalities.

Conclusions: To the authors' knowledge, this is the first study to evidence both active interest and disinterest in direct-to-consumer DNA testing among individuals who are donor conceived. The meanings ascribed to, and uses of, direct-to-consumer DNA testing vary significantly among donor conceived young adults. Findings relating to the relationship between 'informal' and 'formal' information systems, and the absence of guidance and support for those using direct-to-consumer DNA tests, should be considered carefully by practitioners, regulatory bodies and policymakers going forward.

[1] School of Psychology, University of Sussex, Brighton, UK

** Publisher's note: The hyphen in the compound adjective 'donor-conceived' has been removed throughout this article at the request of the author.

*Corresponding author. E-mail address: sophie.zadeh@sussex.ac.uk (S. Zadeh). https://doi.org/10.1016/j.rbmo.2024.103969

Declaration: The author reports no financial or commercial conflicts of interest.

KEY WORDS

Direct-to-consumer DNA testing
Donor
Donor conception
Donor siblings
HFEA
In-depth interviews
Young adults

INTRODUCTION

Direct-to-consumer (DTC) DNA testing is a global phenomenon that is increasing in popularity year on year (*Regalado, 2019*). Provided by companies such as Ancestry and 23andMe, this at-home technology involves providing a saliva sample or cheek swab from which DNA data is abstracted and uploaded to a database of several million users. Along with general curiosity, such tests are taken for different purposes, with users seeking to build family trees, to learn more about their health and genetic predisposition to certain conditions, to learn about their ethnic heritage, and to identify their genetic relations (*Guerrini et al., 2022*; *Horton et al., 2019*; *Nelson, 2008*). The fact that DTC DNA testing may lead to the discovery of unanticipated genetic information, with consequences for mental health outcomes, has also been documented (*Avni et al., 2023*).

Over the last decade, the debate about the implications of DTC DNA testing for donor conception has grown significantly. Scholars have generally argued that DTC DNA testing poses several risks to donor conceived people, their parents and donors, namely that donor conceived people will inadvertently discover their donor conception; parental secrecy of donor conception will be exposed; donor anonymity will be violated; and large same-donor networks will be discovered (*Adams and Allan, 2013*; *Borry et al., 2014*; *Darroch and Smith, 2021*; *Harper et al., 2016*; *Kennett et al., 2019*; *Rotshenker-Olshinka and Dahan, 2020*; *Sadeghi, 2019*; *Woodward, 2015*; *York, 2021*).** Claims about these outcomes are substantiated mainly by anecdotal evidence reported in the media, and by a focus on the context of gamete donation. In terms of the latter, parental secrecy to children about their donor conceived origins was not only the norm historically, but was actively encouraged by those working in the medical profession until relatively recently (*Nordqvist, 2021*; *Richards, 2014*). Secondly, few records were historically kept, and today, few countries have adopted donor identifiability as law, meaning that even among individuals who have been told about their donor conception, little to no information about the donor or others conceived using the same donor (often referred to as 'donor siblings') may be available. Moreover, there is no clear relationship between legally mandated donor identifiability and parental disclosure of donor conception (*Zadeh, 2016*; see also *Lysons et al., 2023*), meaning that even in countries where identifying information about the donor is recognized as the legal right of donor conceived individuals, many will not know that they are donor conceived. Finally, donor disclosure of donation is not common. A 2019 study of over 200 sperm donors showed that a significant proportion had not disclosed their donation to members of their family of origin, and only three participants had told their children (*Graham et al., 2019*). In a more recent study of 396 donors in the USA, just over half of the donors surveyed had disclosed their donation to their families, and approximately one-third of donors with children had shared this information with them (*Wodoslawsky et al., 2023*).

Empirically, little is known about the experiences of DTC DNA testing among those involved in donor conception. Previous research has focused on the experiences of donor conceived people and donors who are part of the UK's Donor Conceived Register (DCR, formerly UK Donor Link). Introduced in 2004, the DCR was the first voluntary register to use DNA as the basis for identifying and making connections among those genetically related through donor conception, and is primarily for those who were conceived or donated prior to 1991, when record keeping relating to donor assisted conception became a legal requirement through the UK Human Fertilisation and Embryology (HFE) Act. This research has focused on why individuals have chosen to use the DCR, as well as the outcomes of doing so (*Frith et al., 2018*; *van den Akker et al., 2015*). More recent research on the Fiom KID-DNA database — a government-funded register in the Netherlands, set up in 2010 — has had a similar focus (*Indekeu et al., 2022a*,*b*). Findings from both bodies of work have shown that test-taking among donor conceived individuals is generally motivated by a desire for information and to make connections, and that outcomes for both donor conceived individuals and donors may vary, from the building of very positive relationships to those that are less positive and, in some cases, negative.

DTC DNA testing differs from these national, voluntary DNA registers in several respects (*Gilman et al., 2024*; *Indekeu et al., 2022b*). Most of all, those who have registered on the DCR or the Fiom KID-DNA database are aware of their donor conception or of having donated (although they may also receive unanticipated information, such as the existence of large same-donor networks; see *Frith et al., 2018*; *Indekeu et al., 2022b*). In contrast, research has shown that those taking DTC DNA tests may learn of their donor conception inadvertently as a result (*Crawshaw, 2018*; *Gilman et al., 2024*; *Guerrini et al., 2022*; *Newton et al., 2023*). The technologies employed by state-funded registers and commercial enterprises also differ, with the former having been shown in some cases to have generated false-positive results (*Adams and Allan, 2013*). Provision for those who do identify a match through the respective databases also differs; concerns have been raised about the lack of information and support provided by commercial testing websites, particularly for those who receive unanticipated results (*Crawshaw, 2018*; *Gilman et al., 2024*; *Indekeu et al., 2022a*, *b)*. Relatedly, the value of support services provided through both the DCR and the Fiom KID-DNA database has been highlighted (*Crawshaw et al., 2016*; *Indekeu et al., 2022b*).

Although limited in scope, the recent empirical research on DTC DNA testing and donor conception offers a valuable insight into the implications of this technology for those affected by donor conception. As above, some of this research has shown that DTC DNA testing may result in individuals learning, for the first time, about their donor conceived origins. A recent community-based survey of 481 donor conceived people found that a significant proportion — approximately one-third of respondents — discovered that they were donor conceived through DTC DNA testing (*We are Donor Conceived, 2020*). Such instances have also been identified in research not focused on the technology (*Bauer and Meier-Credner, 2023*). General research on unexpected discoveries through DTC DNA testing has also compared the experiences of donor conceived people and those who learned of other unanticipated information through this means (*Guerrini et al., 2022*; *Lawton et al., 2023*; *Shepard et al., 2022*). The findings from these studies suggest that donor conceived individuals may be more regretful of using DTC DNA testing, and may experience more negative consequences and emotions, compared with those who have covered other types

of unanticipated information (*Guerrini et al., 2022*; *Lawton et al., 2023*).

Recent studies have captured the experiences of individuals who are aware of donor conception and how it relates to them before taking a DTC DNA test (*Gilman et al., 2024*; *Klotz, 2016*; *Newton et al., 2023*). This research has employed the important social scientific concepts of agency and trust to understand the pursuit of DTC DNA testing by those who are donor conceived, who are often part of 'kinship knowledge regimes' (*Klotz, 2016*) characterized by familial secrecy and clinical gatekeeping. Findings from these studies emphasize the interplay between 'formal' and 'informal' information, with donor conceived individuals using DTC DNA testing to verify and extend the (often limited) 'formal' information given to them, or to gain information to which they have no 'formal' access (*Gilman et al., 2024*; *Newton et al., 2023*). The fact that DTC DNA testing has created new gatekeepers of information — including donor conceived people who match with individuals who are unaware of the relevance of donor conception to them or their family members, and so-called 'DNA detectives', who use their scientific expertise to identify matches between those genetically related through donor conception — has also been highlighted (*Gilman et al., 2024*; *Klotz, 2016*; *Newton et al., 2023*). Scholars have also identified the experience of searching without matching over a period of many years (*Gilman et al., 2024*; *Newton et al., 2023*).

Despite the existing research, many questions about the role of DTC DNA testing in the lives of those affected by donor conception remain. While the analysis by *Gilman et al. (2024)* provides an important insight into the experiences of different parties — donor conceived individuals, their parents, donors and donors' family members — in relation to DTC DNA testing, their research investigated only the experiences of those who had been affected directly by DTC DNA testing. *Newton et al., 2023*, which focused on the experiences of donor conceived people, two-thirds of the 91 survey respondents and an unreported number of individual interviewees had used DTC DNA testing; findings about those who had not taken a test were not reported. Further empirical evidence is therefore needed to understand how individuals who are affected by donor conception — chief among them being those who are donor conceived — make meaning of this technological development, and whether, and to what extent, it plays a role in their lives. The present study sought to address these questions.

MATERIALS AND METHODS

The data reported in this article are taken from a larger research project about donor conceived people in the UK, with a focus on the experiences of those in young adulthood.

Study context

The context for this study, the UK, has seen several legal changes relating to donor conception over the last few decades. The first national legislation on assisted reproduction (and related matters) was introduced in 1990 under the HFE Act. This Act also established the national regulatory body, the Human Fertilisation and Embryology Authority (HFEA), which is responsible for keeping information about donation for those conceived at a UK licensed clinic, including, between 1 August 1991 and 31 March 2005, non-identifying information about the donor, and, on and after 1 April 2005, identifying information about the donor. The HFEA was set up in 1991. For individuals conceived prior to 1 August 1991, non-identifying information may or may not be available from the clinic at which they were conceived; no information is held by the HFEA. The DCR, described in the Introduction, is funded by the HFEA, and is a voluntary register for those wishing to identify genetic connections (including via genetic testing). The DCR is primarily for those conceived before 1 August 1991.

For individuals conceived between 1 August 1991 and 31 March 2005, non-identifying information generally includes a donor's year and country of birth, their physical characteristics, their ethnicity, and whether they had children at the time of donation. It may also include information relating to their occupation and interests, and a brief self-description. Once donor conceived people reach the age of 16 years, they can apply for this information from the HFEA, along with non-identifying information about donor siblings: how many they have, their gender/s, and their year/s of birth. Their parent/s can apply for this non-identifying information at any time after they are born.

At 18 years of age, donor conceived individuals can join Donor Sibling Link (DSL), which allows donor conceived people who are genetically related to one another to exchange contact information. This service is also run by the HFEA. Donors who donated between 1 August 1991 and 31 March 2005 can choose to remove their anonymity retrospectively via the HFEA. Identifying information is only shared with donor conceived individuals at the individual's request.

The legislation following 1 April 2005 does not apply to the individuals in the present study (because of their year of birth).

Sample characteristics

Thirty-three donor conceived adults, aged between 18 and 31 (Mean 24.3, SD 4.31) years, took part in the study. All participants had been conceived using sperm donation at a UK clinic, and their main residence was in the UK. Due to their birth date, most participants ($n = 27$) could request non-identifying information about the donor and those conceived through the same donor from 16 years of age from the HFEA; six participants could not.

As is common in research on donor conception (*Indekeu et al., 2021*), most (76%, $n = 25$) participants in the study were female, and the vast majority (85%, $n = 28$) were White 'of whom most (73%, $n = 24$) described themselves as 'White British'. Further demographic information is reported in TABLE 1. In terms of family composition at conception or birth, participants reported being part of mum and dad (73%, $n = 24$), solo mum (18%, $n = 6$) and two mum (9%, $n = 3$) families. However, a proportion of participants (39%, $n = 13$) across all family types reported that their family composition had changed during childhood due to parental divorce or separation, and/or their parent having met a new partner (including partners of a different gender to partners at the time of conception or birth). In terms of siblings, 17 (52%) participants reported having siblings (not including donor siblings), and 16 (48%) participants were only children. Regarding the age at which participants had been told about their donor conception, 16 (48%) said that they had always known, three (9%) had been told before 10 years of age, four (12%) had been told or found out between 11 and 19 years of age, and 10 (30%) had been told or found out at ≥20 years of age.

TABLE 1 SELF-REPORTED DEMOGRAPHIC CHARACTERISTICS OF THE SAMPLE

Demographic	*n* = 33	%
Gender identity		
Female	25	76
Male	8	24
Sexual orientation		
Straight	22	67
Bisexual	5	15
Gay	2	6
Queer	2	6
Straight questioning	1	3
Prefer not to say	1	3
Ethnicity		
White British	24	73
White/Caucasian	4	12
British	1	3
Mixed race	1	3
English/Irish	1	3
White Caucasian/Jewish	1	3
English/Spanish	1	3
Disability		
None	30	91
Psychological health	1	3
Physical and psychological health	1	3
Learning difficulty	1	3
Relationship status		
Single	13	39
In a relationship	9	27
Cohabiting with a partner	9	27
Married	2	6

Participants were recruited to the study via the UK's largest community networks for donor conception families (Donor Conception Network) and donor conceived people (DCR Registrants' Panel, now Donor Conceived UK) via mailing lists and social media. They were also introduced to the study through snowballing. The original aim was to recruit participants to the study according to family type, acknowledging the literature that has highlighted patterns in individuals' meaning making about donor conception on that basis (*Jadva et al. 2010*). However, given the relative number of individuals of different family backgrounds who wanted to take part, this was not possible. Moreover, given that the family composition of many participants had changed during childhood, sampling according to *a-priori* 'family type' arguably made little sense in practice.

This study received ethical approval from the University College London Institute of Education Research Ethics Committee (Z6364106/2020/01/82, approval date 11 February 2020). It was also approved by the Donor Conception Network Research Ethics Committee. All participants provided written consent to take part in the study. They were each compensated for their time with a £15 gift voucher.

Interviews

Two pilot interviews with donor conceived adults in their 30s were conducted. The feedback from these interviewees and staff at Donor Conception Network resulted in changes to the interview schedule in terms of length, terminology and topics addressed. The final interviews comprised narrative and semi-structured elements. Interviews were based on the timeline method (*de Vries et al., 2017*), which asked participants to plot important experiences to date (both related and unrelated to donor conception) on a timeline, and to reflect upon their past, present and future in relation to their donor conception. Participants were also asked to complete an identity map (*Frost et al., 2020*), and, in so doing, to reflect upon their social roles and identities, and the role of donor conception within these. Semi-structured questions addressed participants' experiences of family life, with peers, within the context of romantic relationships, within communities, and in education. Participants were also asked for their perspectives on legal and social provision relating to donor conception in the UK, their perspectives on and uses of DTC DNA testing, and their interest in and contact with their donor and others conceived using the same donor. Finally, participants were asked about any additional thoughts or experiences that they would like to share for the purposes of the study.

All interviews were conducted over Teams (Microsoft, USA) between November 2020 and September 2021. In terms of length, the interviews lasted for an average of 90 min, with the shortest interview lasting for approximately 43 min and the longest interview lasting for approximately 150 min. Each interview was recorded, transcribed and anonymized during this process. Data collection ended once a saturation point, in terms of interest in the study among prospective participants, seemed to have been reached (within the given time constraints of research funding).

Analysis

The analysis reported in this paper addressed the following research question: what meanings do donor conceived young adults give to DTC DNA testing, and how does DTC DNA testing relate to their lived experiences? To answer this question, reflexive thematic analysis (*Braun and Clarke, 2021*), which makes sense of people's experiences in relation to their social contexts, was used. This analytic approach is in keeping with the study's critical realist orientation (*Willig, 2016*) and its emphasis on personal meaning-making and lived experience, which are understood to be influenced by sociocultural context. The emphasis on personal meaning-making is also in keeping with the study's overall community-engaged approach, described below.

Each interview transcript was coded inductively, line-by-line, with a focus on descriptive codes (e.g. 'DNA test not told parents'; 'DNA test likely to find') rather than identifying latent meanings at this initial stage. Analysis focused on the data in its entirety, rather than answers to questions relating to DTC DNA testing specifically, as participants often discussed DTC DNA testing in response to other parts of the interview (particularly during the timeline exercise). In other words, each participant's interview, in full, served as the unit of analysis. Once initial coding was complete, the codes were refined into broader subthemes. Through the process of writing (*Smart, 2007*) and close engagement with the subthemes and the corresponding data (as per *Braun and Clarke, 2021*), four related themes were identified.

All data collection and analysis were conducted by the author, and the data were analysed separately using a narrative approach, resulting in a depth of familiarity with, and knowledge of, the dataset overall. Given that the author is not donor conceived, it was vital that, in keeping with the recommendations of *Newton (2022)*, the study adopted a community-engaged approach (*Rosenthal, 2016*). In line with this approach and the study's theoretical orientation, during data collection and analysis, the author sought to understand as much as possible about the sociocultural context of donor conception in the UK through knowledge exchange

with staff at Donor Conception Network and volunteers at the DCR Registrants' Panel, with whom the preliminary findings were also discussed. The study benefited overall from a reflexive approach throughout the research process, fostered by knowledge exchange with community members (as above), informal meetings with colleagues working with other data from the study, and by keeping a research journal.

Quotations and terminology

In-depth quotations from participants are reported below; the terminology used in participants' quotations is their own. Each participant is referred to by a number, reflecting the order in which interviews were arranged. Upon first quotation, participants' ages and genders are included.

Where participants discussed individuals conceived using the same donor, they tended to refer to them as 'half-siblings', 'half-brothers' or 'half-sisters', in contrast to siblings with whom they were raised, who were always referred to as 'brothers' and 'sisters' (irrespective of whether or how they were genetically related to them, whether their siblings were also donor conceived, and so on). In a minority of interviews (notably with participants who were only children), participants referred to those conceived using the same donor as 'siblings'. With regards to the donor, some participants referred to their donor as their 'biological father' or 'biological dad', while others used 'donor'.

Where information about genetic and familial relationships was shared with the author and is required for understanding participants' quotations, it is included in brackets.

RESULTS

Of the 33 study participants, 19 (58%) had undertaken at least one DTC DNA test, and 14 (42%) had not. Among those who had used DTC DNA testing, the majority had done so with knowledge of their donor conception; three participants (9%) had learned about their donor conception inadvertently through a DTC DNA test.

In terms of who the participants had identified through DTC DNA testing, five (15%) participants had identified their donor, in some cases with support from other people, including members of communities and search organizations (e.g. DCR Registrants' Panel, Finding Families, Search Angels); and eight (24%) participants had identified at least one donor sibling. One participant (included in the above totals) had identified both their donor and a donor sibling through DNA testing. Four (12%) participants had not identified a match at the time of interview, and three (9%) participants were awaiting their results.

Four related themes explain participants' perspectives and experiences of DTC DNA testing. The first theme, ruptures, captures the sense in which DTC DNA testing either had disturbed or was anticipated to disturb the usual course of individual lives. This theme reflects the ruptures that participants experienced or anticipated for themselves, and the ruptures that DTC DNA testing could create for other people who are donor conceived or otherwise genetically related to the donor. The second theme, disclosures, relates to the ways that participants' interest in or use of DTC DNA testing either had led, or was felt could lead, to disclosure of donor conception, either in their own family, or others' families, including to donor siblings and the donor's family. This theme also captures findings related to participants' disclosure or non-disclosure of their use of DTC DNA testing to their family members. The third theme, webs, explains the ways in which tests, and their results, were described as related to other information, technologies, individuals and families. This theme highlights the connections between different sources of information and different individuals (themselves with different connections to information), and the connections between different individuals that may arise through DTC DNA testing. The fourth theme, temporalities, captures participants' different experiences of DTC DNA testing in real time, and the temporal dimension of their perspectives. This theme emphasizes the various 'speeds' at which DTC DNA testing was experienced, and the different meanings that participants ascribed to test-taking in terms of providing closure versus their open-ended nature. Idiosyncrasies in participants' meaning-making as it relates to these themes are outlined and explained below.

Ruptures

Receiving the results of DTC DNA tests was experienced by some of the participants as a rupture. This was particularly the case for the three participants who had learned of their donor conception through DTC DNA testing. These participants had initially assumed that their results meant that they had uncovered something about their family histories of which they were previously unaware (such as a parent's affair, gamete donation, or adoption):

> On Monday morning I've got this message from someone saying they think they're my half-brother and I'm like, oh, listen, well actually I thought it was probably someone making a terrible mistake, so I was like what an idiot and also [maybe this is] kind of [a] spam thing and then I go on there and see it's got like 26% match or something and I was like oh, bit odd. Looked him up on Facebook, thought oh dear, he's got my face. Spent 2 hours thinking, oh dear, one of my parents may have had an affair.
> (Interviewee 24, 25-year-old female)

> My initial reaction was, I thought that my dad has, er, I thought my dad had donated sperm before he met my mum. And I'd already messaged this girl, and I thought oh God, I've put my foot in it. If my dad doesn't want contact with this girl, I've made contact with her and I shouldn't have done that. So, then I spoke to my, I had to call my dad something like, 'I've put my foot in it, I'm a right mess [laughs] up here'.
> (Interviewee 8, 28-year-old female)

Another participant, who had been conceived using mixed sperm (part of the historical practice of mixing the sperm of a donor and infertile partner), described the process of getting her DTC DNA test results as 'horrific. . . finding out for sure that Dad wasn't Dad' (Interviewee 12, 29-year-old female). The discovery of the absence of genetic connections to other family members was also described as a rupture experienced by both participants and their families. One participant explained that DTC DNA testing was the basis of her brother's acceptance of their genetic (un)relatedness:

> My brother [who is not donor conceived] took it quite badly. He was really nice about it, but he, he wasn't convinced. He said, 'I don't believe it until we get a DNA test'.
> (Interviewee 2, 26-year-old female)

Several other participants, who knew about their donor conception before pursuing DTC DNA testing, described the possibility that testing could lead to a rupture for individuals who either did not

know about their donor conception, or were unaware of their father's donation. Several stated that the commercial organizations providing DTC DNA testing needed to inform users about the possibility of receiving unexpected results that could 'completely unravel your understanding of who you are' (Interviewee 6, 25-year-old male). For some participants, the risk of rupture to others meant that while they had added themselves to the DCR or DSL, they had not pursued DNA testing through any other route:

I certainly, what I wouldn't particularly wanna do is like, ambush people, you know. I think, I would be a bit wary of doing the 23andMe. . . going to someone who I assumed was either my sibling, my donor conceived sibling, or, if my donor had children later in life, you know, sibling, half-sibling via that way. . . What you don't wanna do is, is be like, 'Hello! I'm, I'm, like related to you, 'cause I'm donor conceived, and you are too', and then they go, 'Oh fuck, I didn't know that'. . . you'd want more people to be also mutually kind of interested in, in, to know you. . . which I guess is what I've tried to do on the donor sibling register.

(Interviewee 32, 29-year-old male)

Among those participants who had not pursued DTC DNA testing, some explained that this was because they did not want to identify the donor, did not want to be contacted by the donor or other donor conceived people, or because they were mindful of the possibility of this being a negative experience. These participants were evidently anticipating an experience of rupture, which they sought to avoid, at least for the time being. Some of these participants had also not requested information from the HFEA, believing that this information similarly had the potential to create rupture in their lives, or the lives of their family members:

I think probably in respect of my mum and dad, even though they have said to me if I wanted to get the information they would support me but. . . I don't want to hurt them in a way. . . I know a lot of people have asked about DNA testing, but I think the only reason I haven't done that is for fear that something about my donor will come up and I'm just not ready to do that yet. So yeah, maybe in the future, but at the minute no.

(Interviewee 23, 19-year-old female)

I think one of the last things I'd want is to contacted by someone to say, 'Hello I'm your dad', 'Hello I'm your brother'. Yeah I'm happier, I'm happier living the life I know than trying to live a life, if if I do want to find out then I would like it to be on my terms and I will look and I will deal with that, but I don't really want contact really.

(Interviewee 14, 19-year-old male)

Disclosures

Aligned with the idea that DTC DNA testing may cause a 'rupture' for individuals and families, some participants described how DTC DNA testing had led to the disclosure of donor conception, either in their own family or others' families. Alongside the experiences described above, in which some participants learned of their donor conception through DTC DNA testing, one participant had been told about her donor conception after having expressed an interest in DTC DNA testing to her mother, explaining that, 'She did always plan to tell me, before I had children, but I think I just pushed it along a little bit cos I was looking into DNA' (Interviewee 2). Another participant had his donor conception confirmed by his sister, who had herself learned through a DTC DNA match to a half-brother. A third participant had matched with a half-sister, who did not know she was donor conceived, prompting disclosure in this individual's family:

I never said you are, you know you're my half-sister because that would be sort of, I think for me, that would be, that wouldn't be right. So yeah, I just sort of reached out and then she sort of asked her her mum. . . And her mum kept lying and then finally one day, [half-sister] was like, 'There's something not sitting right here. Er what's wrong?' And then she was like, 'Yeah, yeah she is she is your half-sibling'. Yeah, and that's sort of how that happened. So, I let, I let her parents you know sort of handle that discussion because obviously it's not really my place.

(Interviewee 3, 22-year-old female)

The prospect of disclosing other aspects relating to donor conception, such as the donor's thoughts and feelings about contact, was also mentioned. One participant had identified her donor through DTC DNA testing and the support of a search organization, and had a negative experience, which she felt she would need to explain to any half-siblings:

And me having that weight of knowing the end outcome of — even if I did have a half-sibling like and they were like interested, wanted to know and it was all amazing and like then I'd have to tell them, 'Oh well, I actually know who he is and this was my experience with him like so I don't know how you might be received'. But it's just, it would be that that would play on my mind as well now.

(Interviewee 30, 30-year-old female)

Alongside findings relating to the relational reasons against DTC DNA testing articulated by some of the participants, these findings evoke consideration of the sense of responsibility to others (in these examples, to those conceived using the same donor) that those taking genetic tests may feel (*Hallowell, 1999*). Relatedly, some participants stressed the need for donors to understand the implications of DTC DNA testing — that 'DNA is not anonymous, that's just a fact' (Interviewee 12) — and to thus disclose their donation to any children they might have, and to prepare for contact.

Other participants described not having yet disclosed their decision to use a DTC DNA test to their parents, and, in some cases, struggling to know how to share the information. For some participants, this meant having to ask their matches for discretion with regards to sharing information:

I recently sent off an Ancestry DNA test, and I haven't got the results back yet but I haven't told [my mum] I've done that, I've bought actually 2 kits and one for her and one for me, because I thought if I can eliminate her DNA then I'm left with his, but it's just something that I've never been able to bring myself to say, 'Hey, can we talk about this?'

(Interviewee 9, 29-year-old female)

When I was 16, I, at great expense cos it was very expensive, bought an Ancestry DNA kit, and I got my DNA tested and it turns out my uncle also had his DNA tested so he was like, 'Oh, you're on here!' And I was like, 'Yeah I haven't told my mum yet, so um let's just keep this between us for now'.

(Interviewee 11, 18-year-old female)

Webs

Across the interviews, participants spoke about the connectedness of DTC DNA tests: to information from the HFEA, to other tests, to social media and to other

people. Tests, and their results, were often discussed in terms of their relationship to other information, technologies, individuals and families, suggesting their embeddedness in something like a 'web'.

Several participants emphasized the connectedness of different systems of information. For some participants, taking a DTC DNA test was preceded by accessing further information from the HFEA. For some, this information was what prompted them to use DTC DNA testing, particularly once having received information about others conceived using the same donor: 'I was like, oh right' that's interesting, and then I sent off for a DNA test' (Interviewee 3). For others, obtaining information from the regulator was seen as an opportunity to prepare for the potential results of DNA testing:

If there's 10 half-siblings out there, I'd like to know that [from the HFEA] before I'm sort of in a position you know, before I can see kind of a shock result, oh there's 10 of them and some of them are actually kind of here and you've got a means of contacting them maybe so I don't want to do that until I've had an understanding of what I might discover. . . I'm very curious but I need to take it step by step.

(Interviewee 6)

For a minority of participants, the information held by the HFEA about the donor and/or others conceived using the same donor did not match the information they had received through DNA testing. One participant, who had identified her donor via DTC DNA testing and the help of a search organization, had received inaccurate information about the donor's characteristics (i.e. height, age) from the HFEA. Another participant had identified donor siblings via DTC DNA testing who did not appear on the 'official' list provided to her by the HFEA.

Beyond the relationships described between DTC DNA tests, their results and official sources of information, it was also common for participants engaged in DTC DNA testing to report that they had used multiple DTC DNA testing websites. Some participants reflected upon the fact that the national, voluntary DNA register maintained by the DCR, which was described as having 'only ever matched about one sibling since their existence or something' (Interviewee 32), was less likely to result in matches than commercial sites. It was also common for participants to use social media and search engines alongside DTC DNA testing, once they had received the full name of an individual to whom they had matched. Websites such as Google had enabled some participants to see photographs and videos of their donor, while others had been able to find their donor's address through the search engine. Participants described different reactions to identifying the donor, and the donor's family, in this way:

So, seeing him I was like I can see what characteristics I've got. But learning about him, his family, his kids, it made me a bit upset because, you know, in a different world I could have been raised by him. . . it's just like, you can't just give away sperm and not care about where they end up, and I think that kind of reality hit me when I saw his picture was... you're my dad biologically, but you know, you have no responsibility over me, and that was kind of hurtful.

(Interviewee 22, 24-year-old female)

It does take up a lot of time, particularly like, seeing as there's such an online presence for him, and like, my great-great-great-grandfather. . . I find that really inspiring, I, er, like looking him up.

(Interviewee 19, 31-year-old male)

Although these participants had not yet contacted their donors, others had identified the donor through social media and then contacted him. These participants described themselves with more and less humour as having been 'stalking' (Interviewee 2) or 'harassing' (Interviewee 30) the donor — with the degree of humour in their accounts seemingly relating to their donor's response to contact (positive and negative, respectively).

Some participants described that seeing a photograph of the donor or others conceived through the same donor on social media confirmed for them the results of DTC DNA testing. As one participant said, 'And then I looked her up on Facebook and LinkedIn and I knew' (Interviewee 25, 22-year-old male). For another participant, using social media after DTC DNA testing led her to learn of her donor conception.

So, I was chatting to my best friend about this, 'It's just a bit odd like, I can't work out, and I'll show you the [DNA test] results'. I showed it to people anyway. She was interested in it as well and she did a bit of um, bit of Facebook stalking essentially [laughs] and found this girl on Facebook and all her, uh, statuses were public and I had, I'd done the same I've seen it and looked through it and I was like ok, fine, um, but my friend had done a bit more and she sent me a screen grab of a status that was done and put on Facebook on Father's Day which said something like Happy Father's Day to Donor 12345, that I may never meet, er, with my 30 siblings and blonde hair, blue eyed XYZ, whatever it was.

(Interviewee 8)

In addition to emphasizing the connectedness of DTC DNA testing and other information technologies, participants also described DTC DNA testing in terms of webs of people. For one participant, matching to their donor's cousin was sufficient to identify a match. In other cases, participants had asked their family members to also take a test to 'rule out' certain connections. Other participants described the labour of DTC DNA testing being shared among those conceived using the same donor who were already known to each other (either because of their parents having connected (as in the first quotation below) or through DTC DNA testing (as in the second quotation below):

I guess [half-sibling] doing it, I found out kinda by proxy, and well, you know, I, I, I, yeah, well it would be the same for me, I'm guessing. . . I'm actually quite glad that he just did it. So I didn't have to do it, yeah.

(Interviewee 21, 18-year-old female)

[My half-sibling] went on MyHeritage to try and find, so she's on that one. So if she finds someone on there, obviously she'll tell me about it. But we're gonna do a 23andMe one as well, in case someone comes on there.

(Interviewee 27, 18-year-old female)

Some participants also described a mismatch between those conceived using the same donor in terms of deciding how far to proceed in the search for the donor, and whether to contact him. These participants' accounts highlight that the implications of DTC DNA tests extend beyond the individuals who are taking the test or acting on the resulting information:

I think [half-sister] wants to find out more than me and she wants like a face. But I'm not really too bothered like I suppose like having said [that], he's got children, I think

that'll be interesting, like how old they could be like things 'cause they are probably are a bit older. I think it's just the siblings I'm really bothered about.

(Interviewee 27)

For other participants, the decision to use (or not use) DTC DNA tests was explained as taken jointly by siblings in the same family who had been conceived using the same donor. Findings relating specifically to siblings in the same family highlight the importance of studying their experiences, something that is often overshadowed by the emphasis in the literature on donor siblings (*Bauer, 2022*). These findings also incite further reflection on the relational responsibilities (*Hallowell, 1999*) that arise through such webs, in which the wishes of others affected (or potentially affected) by DTC DNA testing are woven into individual decision-making practices.

Participants also reflected upon the connections to distant relatives they had made through DTC DNA testing. The outcomes they described ranged from developing positive relationships [including to matches with whom they could not easily describe their genetic relationship, as in (She's really lovely. . . a second cousin once removed, her grandma is my great aunt or something' (Interviewee 30)); to choosing not to make contact ['even though I don't speak to any of these second or third cousins, it's just nice seeing them, and being like, wow, like you're related to me' (Interviewee 22)]; to being ignored or unmatched ['This more elderly lady. . .she probably deleted me on Ancestry, so I feel like she possibly knows more. . .' (Interviewee 8)].

Other participants explained that they had no interest in making distant connections, and, as a result, had either limited their engagement with the individuals to whom they had matched through DTC DNA testing, or had decided not to take a test altogether:

I haven't had any, er responses. Except from like third and fourth, third and fourth cousins, and it's like people who are doing genealogy, which I just, I dunno, I can't, I can't think of anything more boring than doing that! [Laughs] I just find it like, absolutely tedious.

(Interviewee 29, 29-year-old male)

I don't even trust these things, that you spit in a thing and it tells you you've got family from the ancient Egyptians or something, you know to me it doesn't really mean anything if you're told you know, if you're told that you come from somewhere, life is about, I think, life is about here now what you're doing, not about [what] 200, 300 years ago someone was doing farming in a field or something you know, I really don't know.

(Interviewee 14)

Temporalities

Several participants described DTC DNA testing in relation to time. For some participants, including those who had not yet identified a match, doing a DTC DNA test was described as having brought them a form of closure. One participant expressly stated that finding her donor and his children was the outcome that she had wanted, without which she would not feel content about her conception. Other participants, however, described feeling a sense of closure once having done a DTC DNA test, regardless of the outcome:

I think now I've reached this point, I'm 18 I've done everything that I can, it's sort of like there's nothing else to look forward to you know, I haven't really got anything else to do, so it's going to, like it's relaxed a little bit, you know I haven't got anywhere else to go to get more information, so I've just accepted it.

(Interviewee 4, 18-year-old female)

I think at the time, at the time when I bought [the test] I was in the head space of I just want answers now, I just want something to change, I think I, I think I was quite unhappy and I thought to myself even if I, even if it yields nothing I can tell myself that I've done it, you know I've gone searching for answers and even if I get nothing I've tried, so it's almost like being proactive to try and check something.

(Interviewee 9)

Another participant, who had not yet done a DTC DNA test, explained that she thought it unlikely that identifying a match would bring her the 'kind of ideal outcome of closure, where you identify all these people, and then the jigsaw fits together. . .I just don't think that it would, personally' (Interviewee 20, 27-year-old female). This participant described herself as not having previously thought about pursuing DTC DNA testing, and, like several others in the study, expressed ambivalence about test-taking in the future, although some had engaged with other means of making connections such as by joining DSL. Some participants explained that the cost of DTC DNA testing had also prohibited them from taking a test for the time being:

I do still toy with it quite frequently but I've never actually sat down and done it, I think probably because it's so expensive.

(Interviewee 5, 23-year-old female)

Sometimes I'm I'm a bit curious to do like a DNA test and maybe I'd get my brother [who is not donor conceived] to do it at the same time so we can see what the differences are a bit. So sometimes I am curious about doing that. Haven't done it yet because it's quite expensive and I'm like maybe I will do it, maybe I won't, not really bothered, and yeah.

(Interviewee 26, 28-year-old female)

In contrast to the sense in which DNA testing was described by some participants as bringing them closure, one participant, who had identified her donor through DTC DNA testing, articulated her experience as follows:

It wasn't like, I got the e-mail and all of a sudden, my life shifted, and I have another family or anything. The way I've felt about it, was a very slow trickle, still ongoing, it's been what, 6 months? Erm [pause] so, it hasn't been as I expected. I did think I would feel, more? I don't know whether because I've still not quite, all the way accepting, even though he has subsequently done an Ancestry DNA test and it does check out, or because I was already in that mindset before of, actually, I'm already okay, as much as I can be with, with where I'm at and, and I can just look forwards now. So I, I wonder if it's because I was already well equipped emotionally? But yeah, it hasn't been as I expected it to be.

(Interviewee 31, 30-year-old female)

For some participants, the possibility of identifying a sibling in the future was felt to be more likely than identifying the donor, due to the donor's assumed interest in maintaining anonymity given the time period in which he had donated, and the fact that he had not re-registered as identifiable. Others expressed that they thought they would be able to identify the donor eventually. One participant reflected on the legal barriers to searching, and described DTC DNA testing as the 'only end option. . .' to identifying the donor and those conceived using the same donor, '. . .which is kind of sad, in a way' (Interviewee 23).

Several participants who had taken a DTC DNA test described the lengthy process of identifying matches. These participants explained that this task was not only time-intensive, but also emotionally laborious, requiring significant emotional effort:

Yes, me and my brother have kind of been doing it together, but we've had to slow down a bit. I think I'm going to try and leave it till this summer because it's quite a lot of work and I find I get a bit obsessed. . . I know we are like on the on the right lines. . . I think it is something that needs a lot of time. . . It's very slow, yeah that's my main experience. It takes a very long time.
(Interviewee 25)

Some participants had purchased DTC DNA tests but either had yet to take them or had waited some time before doing so, either because of the emotional effort involved, or because they wanted to be more prepared, by obtaining information through the HFEA, for the possibility of identifying connections. Other participants, in contrast, described themselves as wanting to identify matches as soon as possible:

[The test] sat on my dresser for you know, over a year, until I was now off work at the minute and I thought, like now or never you know. . .so I did it, I sent it off in the post a few weeks ago, probably just because I had the extra time to think about it.
(Interviewee 9)

The kit was on sale or something so I thought I'll just buy it, I don't have to use it and I haven't done yet because I want to know, get an idea from the HFEA what I might be letting myself in for.
(Interviewee 6)

I'd love to be able to find more, erm, more of my siblings [conceived in other families]. I think when I, um, found out obviously originally, I was thinking I just want to meet them all now? Whereas the others are more like, take a step back, so we're just like, we, one at a time like, they'll come, but I'm, I'd love to just know them all now. And yeah, and know that they're all OK, even if they don't want to talk to me. It just, it feels that we're connected, and I just would like to, to know how they're getting on.
(Interviewee 18, 24-year-old female)

Other participants emphasized the open-ended nature of the tests, and the possibility that matches could emerge at any time. Some participants explained that they were disappointed to have not already identified a match, and that they hoped to do so in the future:

I was actually kind of disappointed with the results because nothing really helpful came back, um I didn't get any matches, I didn't like, no one of my donor's side, um no one immediately stood out like [it] was just disappointing I thought that maybe there was a chance that [pause] I don't know, something would come up. . . I just have like a kind of a vision in my mind of just one day when I'm probably not expecting it maybe to get that e-mail [from a half-sibling] yeah you know, even if we don't end up being best of friends or whatever, being in each other's lives.
(Interviewee 13, 21-year-old female)

A few participants also suggested that individuals in their cohort may be unlikely to test, due to their age and/or to not knowing about their donor conception. These responses tended to be given to questions about the possible events or experiences that might happen in the future, relating to one's donor conception, that would be relevant. The nature of test-taking — including initially waiting for the results, and their open-ended nature — was also described by participants in different ways. The reflections quoted below highlight that not only did participants differ in terms of the pace at which they engaged with DTC DNA testing, but also that their individual feelings of urgency (or lack thereof) were subject to change over time:

I am like, every day, are my results in yet are my results in yet? And I didn't think I'd be like that, but I am 'cause it is quite, like it is an adrenaline like rush. Or is there anyone on there? But then even if there was then [there'd] be like that whole like, oh no, there is someone on there. What do I do? Like do I contact them? . . .You become very blinkered with it, like the whole DNA thing. . . like it does suck you in.
(Interviewee 30)

Every time you get an e-mail saying oh you've got matches and then its, ooh! But then I realized it's just the eighth cousin, so it did raise false hope for quite a while, and I do think that maybe something would come of it, but nothing ever did. . . I used to check like once a week, but I um don't know my password anymore so I don't, I don't check, but I think I am still, I am still a member.
(Interviewee 11)

DISCUSSION

The findings of this study evidence that DTC DNA tests are tools that are inserted into existing familial, personal and legal contexts. They are navigated by donor conceived people with their own lives, histories and temporalities, meaning that the technology is understood and used in different ways by the different individuals who collectively form this group. For some individuals, DTC DNA testing is a means of accessing information that will answer their questions relating to identity and family. For others, DTC DNA testing, and the information that could be learned through it, is not meaningful. Findings therefore reflect previous research that has shown variation among donor conceived individuals with regards to interest in the donor and others conceived using the same donor (*Indekeu et al., 2021*). As such, the findings of this study significantly extend what is already known about the uses of DTC DNA testing among those who are donor conceived (*Gilman et al., 2024*; *Klotz, 2016*; *Newton et al., 2023*) by highlighting that donor conceived individuals may or may not be interested in using this technology. In other words, how individuals understand DTC DNA testing — and, for those who test, their results — depends on the meanings they ascribe to donor conception and the ways they understand themselves in relation to the donor and others conceived through the same donor (see also *Newton et al., 2023*). For those who learn about their donor conception through DTC DNA testing, this process of meaning-making is *ad hoc*, and often relies upon the initial exclusion of more familiar family narratives (e.g. the existence of an affair; *Smart, 2011*).

Despite the numerous articles that have foregrounded the possibilities that DTC DNA testing will result in donor conceived people inadvertently discovering their donor conception, in parental secrecy being exposed, and in donor anonymity being violated, the findings of this study suggest that experiences of DTC DNA testing among those who are donor conceived are altogether more nuanced. As in other research on this topic (*Crawshaw, 2018*; *Gilman et al., 2024*; *Newton et al., 2023*), discovery of donor conception through DTC DNA testing was a reality for three of the participants in this study. For the majority who chose to use DTC DNA testing, test-taking came after parental disclosure (including what has been termed 'later' disclosure (*Lampic*

et al., 2022)). The results also suggest that DTC DNA testing may itself produce new forms of (non)disclosure (see also Gilman et al., 2024; Newton et al., 2023), with donor conceived individuals who have taken tests being charged with decision-making about sharing this information with their parents, and with other people to whom they have matched, including matches who are themselves unaware of their donor conception. As such, donor conceived individuals using DTC DNA tests may be said to be involved in new forms of genetic responsibility (Hallowell, 1999), through which they take on the management of their own, and others', genetic information, becoming (in some cases) accidental gatekeepers of such information. Extending existing conceptual efforts in the field (Gilman, 2022), more social scientific attention should now be given to the relational responsibilities that arise from DTC DNA testing, and how these are managed by donor conceived individuals who are aware of their conception (e.g. through testing/not testing, disclosing/not disclosing test-taking to family members, disclosing/not disclosing donor conception to matches, and so on). Unlike state-funded, voluntary DNA registers, which have provided valuable support for those making connections, including intermediary services for contact (Crawshaw et al., 2016; Indekeu et al., 2022b), it is worth emphasizing that commercial companies providing DTC DNA tests offer no such provision. This is particularly concerning given the finding that some donor conceived individuals may be more likely to engage with DTC DNA testing than they are with state-managed services, believing the former to be more popular and more effective in terms of matching than the latter.

The findings relating to the relative status of the information gained through DTC DNA testing and other sources of information are important for several reasons. Firstly, as in other research (Gilman et al., 2024; Newton et al., 2023), this study found that many donor conceived individuals used DTC DNA testing to circumvent existing legislation, which, given their date of birth, provides them with limited or no access to information about the donor and others conceived through the same donor. However, while the accounts of some of the donor conceived people in this study seem to suggest that test-taking is a means of asserting agency (see also Klotz, 2016; Newton et al., 2023), it is also clear that the need to engage with DTC DNA testing is perceived by some as the only option, given inadequate access to 'official' information. In these circumstances, the language of agency seems unsatisfactory. The fact that some of the donor conceived people in this study had also received 'official' information that differed from the information provided by DTC DNA testing is worrying. This finding should be particularly concerning to the national regulator, given the large-scale plans to 'open the register' for those with access to identifiable donor information who are coming of age in 2023 and beyond.

This study also sheds further light on previous research that has interpreted DTC DNA testing in terms of its relationship to new modes of 'digital sociality' (Newton et al., 2023). In addition to the role of online communities and networks, findings highlight the use of multiple DTC DNA tests, and the use of DTC DNA testing alongside social media and search engines, among those who are donor conceived. These avenues have meant that some of the donor conceived people in this study were able to see a photograph or watch a video of the donor without contacting him, to find the donor and contact him via Facebook, and, in one case, to obtain the donor's mailing address. These examples seem to confirm that DTC DNA testing may, in some cases, result in donor non-anonymity (Harper et al., 2016). Findings also make clear the role of siblings within the same family in experiences of searching, a theme that itself requires further exploration (Bauer, 2022). Altogether, the results highlight that DTC DNA testing is part of a whole host of technological and relational resources that may be drawn upon in the process of identifying donor connections. Best practice guidelines with regards to DTC DNA testing (ESHRE Working Group on Reproductive Donation et al., 2022) should be updated to reflect these possibilities more fully.

In addition to identifying the different engagements of donor conceived people with DTC DNA testing, the findings of this study also highlight their different expectations of DTC DNA testing. Examples include individuals who were concerned they would uncover undesirable information and therefore chose not to test; those who were testing for general genealogical curiosity and had matched with donor connections inadvertently; those who anticipated finding a match and were disappointed not to have done so; those who did not anticipate finding a match and were surprised to have done so; and those who had ultimately identified the donor and had either a positive or negative experience that was contrary to their expectations. These different expectations and experiences complicate the idea that DTC DNA testing will result in specific outcomes for those who are donor conceived. While scholars have emphasized the risk that DTC DNA testing may result in the identification of matches who have different expectations of contact (Pennings, 2019; Zadeh, 2016), the role of expectations in shaping the experiences of DTC DNA testing itself warrants further consideration. To date, very few studies have addressed this (for an exception, see the work by van den Akker et al. (2015) on the UK voluntary register). Given what is known about the relationship between unanticipated DTC DNA test outcomes and mental health outcomes (Avni et al., 2023), further research on this topic should be conducted.

Finally, the findings highlight the various speeds with which donor conceived people may engage with DTC DNA tests, including long periods of reflection prior to test-taking (van den Akker et al., 2015), and the perceived open-ended nature of test outcomes, which means that individuals may be waiting for an indefinite period before identifying a match (Gilman et al., 2024; Newton et al., 2023). In contrast to the academic debate on this topic, which has tended to emphasize the immediacy of the risks posed by DTC DNA testing, attention to the temporalities involved in DTC DNA testing reveals the complexities of the technology and how it is both experienced and managed by those who are donor conceived. Given these findings, those writing about the potentialities of DTC DNA testing should be mindful not to overstate possible outcomes, and/or to make claims about their likelihood (e.g. likelihood of making connections, see Harper et al. (2016) and Kennett et al. (2019)), going forward.

The limitations of this study include the homogeneity of participants in terms of their sociodemographic characteristics, which means that the study's key finding – that meanings ascribed to, and uses of, DTC DNA testing vary significantly among

donor conceived young adults — cannot be subject to further scrutiny in relation to factors such as gender, which has previously been shown to be associated with individuals' level of interest in donor siblings (with women expressing more interest than men; see *Indekeu et al. (2021)*). Whether such trends are also true of users of DTC DNA testing would be valuable to research. Future community-engaged research on this topic, particularly idiographic research, could also involve participant consultation in the analytic process (*Rosenthal, 2016*), thus further overcoming the possible limitations of researcher positionality. In the meantime, these findings make it clear that DTC DNA testing is understood and used in different ways by different donor conceived young adults. The academic literature on this topic should therefore avoid presenting a 'single story' (*Teman, 2019*) about the phenomenon that cannot be substantiated empirically.

DATA AVAILABILITY

The data that has been used is confidential.

ACKNOWLEDGEMENTS

The author wishes to thank all those who took part in the study and distributed it to their networks. In addition, the author wishes to thank the staff at Donor Conception Network and Freddie Howell (Former Chair of the DCR Registrants' Panel) for kindly providing their time and expertise during project meetings, the staff at Donor Conception Network for providing expert feedback on the interview schedule, and both Yael Ilan-Clarke (Donor Conception Network) and Freddie Howell for providing assistance with recruitment. The support of the UK Economic and Social Research Council (New Investigator Award ES/S015426/1) is gratefully acknowledged.

REFERENCES

Adams, D., Allan, S., 2013. Building a family tree: Donor-conceived people, DNA tracing and donor 'anonymity. Australian Journal of Adoption 7 (2), 1–16.

Avni, C., Sinai, D., Blasbalg, U., Toren, P., 2023. Discovering your presumed father is not your biological father: Psychiatric ramifications of independently uncovered non-paternity events resulting from direct-to-consumer DNA testing. Psychiatry Research 323. https://doi.org/10.1016/j.psychres.2023.115142.

Bauer, T., 2022. Finding out about being donor-conceived, social siblings, and the management of (non-)knowing. Journal of Philosophy and Ethics in Health Care and Medicine 16, 49–63.

Bauer, T., Meier-Credner, A., 2023. Circumstances leading to finding out about being donor-conceived and its perceived impact on family relationships: A survey of adults conceived via anonymous donor insemination in Germany. Social Sciences 12 (3), 155. https://doi.org/10.3390/socsci12030155.

Borry, P., Rusu, O., Dondorp, W., De Wert, G., Knoppers, B.M., Howard, H.C., 2014. Anonymity 2.0: Direct-to-consumer genetic testing and donor conception. Fertility and Sterility 101 (3), 630–632.

Braun, V., Clarke, V., 2021. Thematic analysis: A practical guide. SAGE Publications, California.

Crawshaw, M., Frith, L., van den Akker, O., Blyth, E., 2016. Voluntary DNA-based information exchange and contact services following donor conception: an analysis of service users' needs. New Genetics and Society 35 (4), 372–392.

Crawshaw, M., 2018. Direct-to-consumer DNA testing: the fallout for individuals and their families unexpectedly learning of their donor conception origins. Human Fertility 21 (4), 225–228.

Darroch, F., Smith, I., 2021. Establishing identity: How direct-to-consumer genetic testing challenges the assumption of donor anonymity. Family Court Review 59 (1), 103–120.

de Vries, B., LeBlanc, A.J., Frost, D.M., Alston-Stepnitz, E., Stephenson, R., Woodyatt, C., 2017. The relationship timeline: A method for the study of shared lived experiences in relational contexts. Advances in Life Course Research 32, 55–64.

ESHRE Working Group on Reproductive Donation, Kirkman-Brown, J., Calhaz-Jorge, C., Dancet, E.A.F., Lundin, K., Martins, M., Tilleman, K., Thorn, P., Vermeulen, N., Frith, L., 2022. Good practice recommendations for information provision for those involved in reproductive donation. Human Reproduction Open 2022 (1) hoac001.

Frith, L., Blyth, E., Crawshaw, M., van den Akker, O., 2018. Searching for 'relations' using a DNA linking register by adults conceived following sperm donation. BioSocieties 13, 170–189.

Frost, D., Hammack, P.L., Wilson, B.D.M., Russell, S., Lightfoot, M., Meyer, I.H., 2020. The qualitative interview in psychology and the study of social change: Sexual identity development, minority stress, and health in the generations study. Qualitative Psychology 7 (3), 245–266.

Gilman, L., 2022. The 'selfish' element: How sperm and egg donors construct plausibly moral accounts of the decision to donate. Sociology 56 (2), 227–243.

Gilman, L., Redhead, C., Hudson, N., Fox, M., Nordqvist, P., MacCallum, F., Kirkman-Brown, J., Frith, L., 2024. Direct-to-consumer genetic testing and the changing landscape of gamete donor conception: Key issues for practitioners and stakeholders. Reproductive BioMedicine Online 48 (1), 103421.

Graham, S., Freeman, T., Jadva, V., 2019. A comparison of the characteristics, motivations, preferences and expectations of men donating sperm online or through a sperm bank. Human Reproduction 34 (11), 2208–2218.

Guerrini, C.J., Robinson, J.O., Bloss, C.C., Brooks, W.B., Fullerton, S.M., Kirkpatrick, B., Lee, S.S., Majumder, M., Pereira, S., Schuman, O., McGuire, A.L., 2022. Family secrets: Experiences and outcomes of participating in direct-to-consumer genetic relative-finder services. American Journal of Human Genetics 109 (3), 486–497.

Hallowell, N., 1999. Doing the right thing: genetic risk and responsibility. Sociology of Health and Illness 21 (5), 597–621.

Harper, J.C., Kennett, D., Reisel, D., 2016. The end of donor anonymity: how genetic testing is likely to drive anonymous gamete donation out of business. Human Reproduction 31 (6), 1135–1140.

Horton, R., Crawford, G., Freeman, L., Fenwick, A., Wright, C.F., Lucassen, A., 2019. Direct-to-consumer DNA testing. British Medical Journal 367, l5688.

Indekeu, A., Maas, A.J.B.M., McCormick, E., Benward, J., Scheib, J.E., 2021. Factors associated with searching for people related through donor conception among donor-conceived people, parents, and donors: A systematic review. Fertility and Sterility Reviews 2 (2), 93–119.

Indekeu, A., Bolt, S.H., Maas, A.J.B.M., 2022a. Meeting multiple same-donor offspring: Psychosocial challenges. Human Fertility 25 (4), 677–687.

Indekeu, A., Prinsen, C.F.M., Maas, A.J.B.M., 2022b. Lessons from 10 years' experience running the Fiom KID-DNA database, a voluntary DNA-linking register for donor-conceived people and donors in The Netherlands. Human Fertility. https://doi.org/10.1080/14647273.2022.2144772.

Jadva, V., Freeman, T., Kramer, W., Golombok, S., 2010. Experiences of offspring searching for and contacting their donor siblings and donor. Reproductive BioMedicine Online 20 (4), 523–532.

Kennett, D., Reisel, D., Harper, J., 2019. Genetic databases and donor anonymity. Human Reproduction 34 (9), 1848–1849.

Klotz, M., 2016. Wayward relations: Novel searches of the donor-conceived for genetic kinship. Medical Anthropology 35 (1), 45–57.

Lampic, C., Skoog Svanberg, A., Gudmundsson, J., Leandersson, P., Solensten, N.G., Thurin-Kjellberg, A., Wånggren, K., Sydsjö, G., 2022. National survey of donor-conceived individuals who requested information about their sperm donor-experiences from 17 years of identity releases in Sweden. Human Reproduction 37 (3), 510–521.

Lawton, B.L., Pyott, L.C., Deyerin, K.R., Foeman, A.K., 2023. Experiences of misattributed parentage communities: Impacts of discovering new familial kinships. Journal of Family History. https://doi.org/10.1177/03631990231156176.

Lysons, J., Imrie, S., Jadva, V., Golombok, S., 2023. Families created via identity-release egg donation: Disclosure and an exploration of donor threat in early childhood. Reproductive

BioMedicine Online 47 (4). https://doi.org/10.1016/j.rbmo.2023.05.007.
Nelson, A., 2008. Bio science: Genetic genealogy testing and the pursuit of African ancestry. Social Studies of Science 38 (5), 759–783.
Newton, G, 2022. Doing reflexivity in research on donor conception: Examining moments of bonding and becoming. In: Shaw, R.M. (Ed.), Reproductive citizenship: Technologies, rights and relationships. Palgrave Macmillan, Singapore, pp. 279–301.
Newton, G., Drysdale, K., Zappavigna, M., Newman, C.E., 2023. Truth, proof, sleuth: Trust in direct-to-consumer DNA testing and other sources of identity information among Australian donor-conceived people. Sociology 57 (1), 36–53.
Nordqvist, P., 2021. Telling reproductive stories: Social scripts, relationality and donor conception. Sociology 55 (4), 677–695.
Pennings, G., 2019. Genetic databases and the future of donor anonymity. Human Reproduction 34 (5), 786–790.
Regalado, A. (2019). More than 26 million people have taken an at-home ancestry test. *MIT Technology Review*. Available from: https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/.
Richards, M., 2014. A British history of collaborative reproduction and the rise of the genetic connection. In: Freeman, T., Graham, S., Ebtehaj, F., Richards, M. (Eds.), Relatedness in Assisted Reproduction: Families, Origins and Identities. CUP, Cambridge, pp. 21–43.
Rosenthal, L., 2016. Incorporating intersectionality into psychology: An opportunity to promote social justice and equity. American Psychologist 71 (6), 474–485.
Rotshenker-Olshinka, K., Dahan, M.H., 2020. Fertility care in the era of commercial direct-to-consumer home DNA kits: issues to ponder. Journal of Assisted Reproduction and Genetics 37, 689–692.
Sadeghi, M.R., 2019. Coming soon: Disclosing the identity of donors by genealogical tests of donor offspring [Editorial]. Journal of Reproduction and Infertility 20 (3), 119–120.
Shepard, A., Diamond, D., Willard, L., Staples, J., Martin, K., Witherspoon, N., 2022. Discovering misattributed paternity after DNA testing and its impact on psychological well-being and identity formation. American Journal of Qualitative Research 6 (3), 189–211.
Smart, C., 2007. Personal Life: New Directions in Sociological Thinking. Polity Press, Cambridge.
Smart, C., 2011. Families, Secrets and Memories. Sociology 45 (4), 539–553.
Teman, E., 2019. The power of the single story: Surrogacy and social media in Israel. Medical Anthropology 38 (3), 282–294.
Van den Akker, O.B., Crawshaw, M.A., Blyth, E.D., Frith, L.J., 2015. Expectations and experiences of gamete donors and donor-conceived adults searching for genetic relatives using DNA linking through a voluntary register. Human Reproduction 30 (1), 111–121.
We Are Donor Conceived (2020). *2020 We are Donor Conceived survey report*. Available from: https://www.wearedonorconceived.com/2020-survey-top/2020-we-are-donor-conceived-survey/
Willig, C., 2016. Constructivism and 'the real world': Can they coexist? QMiP Bulletin 21, Spring 2016.
Wodoslawsky, S., Fatunbi, J., Mercier, R., Braverman, A.M., 2023. Sperm donor attitudes and experiences with direct-to-consumer DNA testing. Fertility and Sterility Reports 4 (1), 36–42.
Woodward, J.T., 2015. Third-party reproduction in the internet age: The new, patient-centred landscape. Fertility and Sterility 104 (3), 525–530.
York, M.C., 2021. I just took a DNA test —turns out I'm 100% breaching my donor anonymity contract: direct-to-consumer DNA testing and parental medical-decision-making. Indiana Journal of Global Legal Studies 28 (2), 293–326.
Zadeh, S., 2016. Disclosure of donor conception in the era of non-anonymity: Safeguarding and promoting the interests of donor-conceived individuals? Human Reproduction 31 (11), 2416–2420.

Received 24 October 2023; received in revised form 13 March 2024; accepted 18 March 2024.

RBMO

ELSEVIER

ARTICLE

How to talk to young adults about fertility

BIOGRAPHY

Randi Sylvest is a Postdoc, PhD and MSc in Public Health from Copenhagen University, Denmark. She undertakes qualitative and quantitative research with a focus on men, family building, fertility treatment and fertility awareness, and has conducted over 100 interviews. She has authored over 20 original papers with peer review and presented at several big conferences worldwide.

Randi Sylvest[1,2,*], Emily Koert[2], Julie Balslev[2], Michala Steenberg[2], Lone Schmidt[2], Søren Ziebe[1]

KEY MESSAGE

This study contributes to the understanding and implementation of educational interventions on future fertility awareness targeted to and effective among young adults. Information needs to be targeted to the audience and provided at an appropriate stage in the lifespan, starting with planting the seed when they are young.

ABSTRACT

Research question: How knowledgeable are Danish young adults about fertility and what are their attitudes towards learning about their reproductive biology?

Design: The study was conducted at different educational institutions with 11 focus-group discussions that included a total of 47 participants (aged 18–29 years). Qualitative content analysis was used. The participants' fertility knowledge score was measured using the Cardiff Fertility Knowledge Scale.

Results: The participants had an overall fertility knowledge score of 54%. Focus-group data showed that they thought it was important to learn about fertility and how to protect their fertility potential regardless of whether or not they wanted children. Providing knowledge is like planting a seed in the young adults. They wanted to hear about fertility in multifaceted ways and formats, and believed the information should be delivered by professionals, but developed in partnership with young people. The double-edged sword of knowledge and the consequence of knowledge made them hesitant or less open to learning.

Conclusions: Recommendations from this study are to tailor fertility information to young people, with due cognisance of their developmental stage, and ideally from an earlier age.

INTRODUCTION

Over the past 30 years women and men in many countries have increasingly been postponing parenthood. The average age at first birth for Danish men and women is the highest ever recorded, at 31.6 and 29.9 years, respectively (*Statistics Denmark, 2023*). Advanced female age and advanced male age are associated with lower fertility and a higher risk of adverse birth outcomes (*Brandt et al., 2019*). A recent global review study from World Health Organization showed an average lifetime prevalence of infertility of 17.5%, with regional differences (*Cox et al., 2022*). In Denmark, a previous population-based study reported a lifetime infertility prevalence of 26% among women attempting to conceive (*Schmidt et al., 1995*).

Many factors, including sociocultural, political and biological, influence decisions about whether, when and how to conceive (*Bodin et al., 2021a*). Several review studies have shown that women and men tend to underestimate the age-related decline in fecundity and overestimate the success rate of medically assisted reproductive technologies (*Delbaere et al., 2021*; *Hammarberg et al., 2017*; *Pedro et al., 2018*; *Ren et al., 2023*). Postponement of

[1] Department of Gynecology, Fertility and Obstetrics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
[2] Department of Public Health, University of Copenhagen, Denmark

*Corresponding author. E-mail address: Randi.lykke-sylvest@regionh.dk (R. Sylvest). https://doi.org/10.1016/j.rbmo.2024.103937 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEY WORDS

Education
Fertility awareness
Fertility knowledge
Qualitative
Youth

parenthood may be a consequence of insufficient knowledge about reproductive biology and risk factors affecting fertility. Vassard and colleagues demonstrated that participants who overestimated the probability of pregnancy in a 30-year-old woman were more inclined to attempt or desire their first child at a later age (*Vassard et al., 2016*). Furthermore, the misperception that assisted reproduction can compensate for difficulties in conceiving at an advanced reproductive age may lead to a postponement of parenthood (*Daniluk & Koert, 2013*; *Hammarberg et al., 2017*; *Pedro et al., 2018*).

Fertility awareness is defined as:

The understanding of reproduction, fecundity, fecundability, and related individual risk factors (e.g. advanced age, sexual health factors such as sexually transmitted infections, and life style factors such as smoking, obesity) and non-individual risk factors (e.g. environmental and work place factors); including the awareness of societal and cultural factors affecting options to meet reproductive family planning, as well as family building needs. (*Zegers-Hochschild et al., 2017*: 8)

Fertility awareness interventions take a preventive focus, with the goal of reducing future infertility and promoting informed decision making so that people can meet their family-building goals.

Improving fertility knowledge and awareness continues to be a crucial component of public health initiatives aiming at preventing involuntary childlessness and achieving desired family-building intentions. People often change their perspectives on family building at different stages in life; therefore, fertility-awareness strategies need to be tailored differently to suit the target population in different life phases. In their mixed method study, Grace and colleagues found six key categories of family building preferences – Avoiders, Betweeners, Completers, Desirers, Expectants and Flexers – for whom fertility education strategies could be tailored differently to suit their situation and readiness (*Grace et al., 2022*).

A growing body of research suggests that young adults want to know more about fertility, preferably as part of their secondary education. Ragnar and co-workers found that young adults wish to receive fertility information through a variety of sources and at several time points (*Ragnar et al., 2018*). Likewise, Bodin and collaborators found that the youngest participants (age 17–24 years) were the most receptive to fertility information, and that young adults (age 25–35 years) were less receptive; however, they said that they wished they had learned more about fertility in school (*Bodin et al., 2021a*). A Danish study found that young men wanted general fertility information to be a mandatory part of the education curriculum in primary and secondary schools (*Berthelsen et al., 2021*).

However, if fertility educational initiatives are going to be effective, there is a need for an understanding of and interest in the target population (*Garcia et al., 2016*). Qualitative methods are often an optimal methodology to elicit data on the feasibility and acceptability of new interventions and processes (*Bowen et al., 2009*). Thus, the aim of this study was to explore the knowledge of Danish young adults about fertility, their attitudes towards learning about reproductive biology and how they want to receive this information.

MATERIALS AND METHODS

Recruitment

This study adopts a qualitative approach involving focus groups with young adults aged 18–29 years. Eligible participants came from both rural areas and larger cities in Denmark. Institutions willing to participate received information about the study through an invitation letter.

Ethics and data protection

The study followed the principles of the Declaration of Helsinki II for medical research. According to Danish legislation, interview studies do not require approval from a scientific ethics committee. Participants were provided with written informed consent to participate in the study, which they filled out and signed before the interview started. In the introduction to each focus group, the participants received information about the purpose of the study and their rights as research participants to withdraw at any time and to maintain their confidentiality. All collected data (transcripts, recordings, signed informed consent forms and Cardiff Fertility Knowledge Scale answers) were stored according to Danish Data Protection Agency rules regarding data protection.

Sampling

A purposive sampling with maximum variation was used regarding educational level, since family-building aspirations could be influenced by different levels of education (*Aguinaldo et al., 2014*). The participants were recruited from preparatory basic education, high school, Bachelor-level education and technical school.

Data collection

A semi-structured interview guide was used in the 11 focus groups. At the beginning of the focus group discussion, the participants received a short questionnaire to measure their knowledge relating to fertility. The Cardiff Fertility Knowledge Scale (*Bunting et al., 2013*) was used. This validated scale assesses fertility knowledge using a 13-item true/false scale concerned with risk factors, misconceptions and basic fertility facts on, for example, smoking, history of sexually transmitted infections, and whether a woman is fertile even without menstrual periods. A correct answer was assigned 1 point and an incorrect or do not know answer assigned 0 points. Points were summed, divided by the total number of questions and multiplied by 100 to produce a percentage correct fertility knowledge score with a range of 0% to 100%. In the original Cardiff Fertility Knowledge Scale, the internal consistency coefficient alpha (Cronbach's α) was 0.79 (*Bunting et al., 2013*).

The participants were also shown a fertility awareness poster with nine facts about fertility and reproduction during the interview. The poster was used as a tool to provoke discussion in the focus group discussion. The fertility awareness poster was developed by the International Fertility Education Initiative (later renamed the International Reproductive Health Collaboration), within the European Society of Human Reproductive and Embryology. The poster can be found at https://fertilityeurope.eu/fertility-education-posters/.

Data collection took place in spring 2023. In qualitative research, data saturation is used to ensure the trustworthiness of the data collection and analysis. Data saturation requires that data are collected until no new themes or information appears from each focus group (*Saunders et al., 2018*). In this study, data saturation was achieved after 11 focus group discussions.

Interview guide

The interview guide consisted of five topics. The first topic included questions regarding the participants' thoughts about the Cardiff Fertility Knowledge Scale after they had filled out the scale. The second topic was their initial thoughts about their own family building in the future. The third topic included questions concerning the participants' thoughts about the poster and whether or not the poster, as well as fertility information in general, was relevant to them according to their life situation. The fourth topic regarded their intentions to act upon the new knowledge. The fifth topic concerned questions about other fertility interventions in Denmark, the participants' needs and desires for fertility information and how they wanted fertility information to be disseminated in the future.

The interview guide was prepared in collaboration with the co-authors. The questions were open-ended and exploratory. A pilot interview with three participants was conducted before the first focus group discussion and minor edits were made to the interview guide. The focus groups varied between four and six participants per group. The focus group discussions were facilitated by the first author (R.S.) who conducted the focus group discussions as moderator. R.S. had conducted over 100 interviews and 20 focus group discussions before the current study. For some of the current focus group discussions, an observer was also present (J.B. or M.S.).

The interviews lasted on average 90 min (range 60–120 min). The interviews were in Danish and the quotes were translated into English by a native English speaker (E. K.) who is also able to understand Danish. Field notes were made during and after the focus groups.

Analysis

Using the Cardiff Fertility Knowledge Scale (*Bunting et al., 2013*), the percentage correct fertility knowledge score, with a range of 0% to 100%, was calculated for the 47 participants. Descriptive statistics were used to present the results. A chi-squared test was used to compare fertility knowledge scores. Statistical significance was defined as a two-sided *P*-value of <0.05. Descriptive statistics were generated using the statistical software SAS Entreprise Guide 8.3.

The focus group discussions were recorded, anonymized and transcribed in full. Participants were anonymized in the transcripts. Data were analysed using qualitative content analysis following the method of *Graneheim and Lundman (2004)*. Interview transcripts were first read carefully to develop a sense of the content. Salient sentences and paragraphs that related to the study questions were identified and labelled with a code reflecting their meaning. Codes were grouped into categories and sub–themes, and finally overall themes were identified based on the similarities and differences. Consensus was achieved through discussions, and changes were made to the codes, sub-themes and themes at each stage after discussion with the co-authors. To increase trustworthiness in the reporting of the study, the analysis followed the COnsolidated criteria for REporting Qualitative research (COREQ) (*Tong et al., 2007*).

RESULTS

This study is a qualitative study of 11 focus groups (mixed gender) with a total of 47 young adults. Fifteen different public educational institutions were contacted, of which 11 agreed to participate. The participants were between 18 and 29 years old and were recruited from different types of public educational institution (preparatory basic education [5 students], technical schools [7 students], high schools [18 students] and professional Bachelor's education [17 students]). The study participants were single or cohabiting young adults living throughout Denmark. The majority were female (36 women and 11 men). Most of the participants (83%) wanted to have children in the future. Only two of the participants already had children.

Fertility knowledge – the Cardiff Fertility Knowledge Scale

All the participants answered all the questions on the Cardiff Fertility Knowledge Scale. The overall mean fertility knowledge score was 54% (13–93%). There was no significant difference between genders ($P = 0.9$). The men had a fertility knowledge score of 55%, and the women a score of 54%. No significant difference was found according to the type of school ($P = 0.3$). The participants from high schools had a fertility knowledge score of 55%, professional Bachelor's students had a score of 54%, students from technical schools had a score of 51%, and students from preparatory basic education had a score of 49%.

It was noted that many participants had little knowledge surrounding fertility problems and how to protect their fertility. Most of the participants knew that female fertility declines rapidly (67%), but only a third of the participants knew that increased age in males also leads to decreased fertility (33%).

Qualitative results

The following themes were found during analysis: importance of fertility knowledge; sensitivity to, and respect for, the personal nature of fertility and family-building choices; planting a seed; how, what and from whom; the double-edged sword of knowledge/consequence of knowledge; and readiness and priorities (FIGURE 1).

Importance of fertility knowledge

The participants believed that it was important to learn about fertility regardless of whether or not they wanted children, and they thought it should be a mandatory part of the curriculum during their education:

> Really important that we are informed much more about this, because I don't think there are many people who just walk around and know about this. (Male, 18 years, high school)

> Nice to be able to make some informed choices. (Female, 22 years, technical school)

> When you are 18 years old, you are perhaps a little more aware in your head about how it all works. (Female, 24 years, Bachelor's education)

The participants thought it was important to make it normal to discuss the topics of sex and fertility. Some of the participants thought their earlier education (if any) had been lacking in information about fertility:

> I have never been taught anything other than how not to have children. I think it has done something to my head. (Male, 26 years, technical school)

Even though some of the participants did not want children in the future, they still wanted to know about fertility in order to help others, protect themselves from unwanted pregnancies or have the

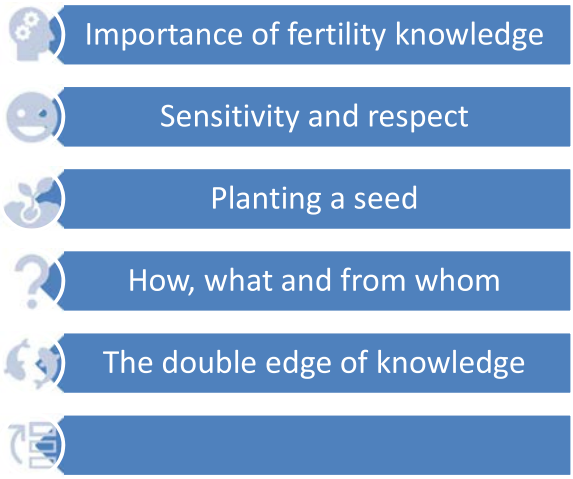


FIGURE 1 Analytical themes identified in the study.

knowledge if they changed their intentions about family in the future.

Most of the participants were uncertain about their fertility knowledge. They were not aware of what they were supposed to know, due to an absence of education. Some became surprised about their lack of knowledge during the interview:

I think there are some questions we should know the answer to now that we are getting older, to which I do not know the answer. (Female, 18 years, high school)

Sensitivity to, and respect for, the personal nature of fertility and family building choices

Many of the participants thought that having children is a private and personal issue and they do not want others to have opinions about it:

I think it is difficult for me to see that others should have something to say about when I should have children. I agree that it's best to have children when you're younger, etc., at least easiest, but I think it's very personal and an individual thing. (Male, 19 years, high school)

The participants wanted to be taught about fertility, but they did not want to feel as though educators were preaching to them. Many of the participants expressed that 'finger-wagging' about what to do to protect their fertility was not an effective approach when dealing with this population:

That kindergarten mentality. 'You say I can't, so I do', and we will probably never let go of that. It will always be there that rebel. (Female, 20 years, Bachelor's education)

They saw a clear boundary between talking about fertility and trying to push them to have children, and it was important for educational efforts to not cross this line:

I think I might be a bit uncomfortable if you guys come over, because we're planning to have kids – like, I don't want you trying to persuade us to have them. There needs to be a clear line that you're not trying to convince them, just informing them. (Female, 19 years, high school)

Some of the participants interpreted the received information into a narrative of pressure rather than education and improved knowledge.

Planting a seed

The participants wanted to learn about fertility at an early age, so they would have the relevant information prior to starting their family in the future. It is like planting a seed in the young adults:

I stored it somehow subconsciously without having to use it right now. (Female, 29 years, Bachelor's education)

Furthermore, they believed the information should be repeated in order for it to be incorporated into their knowledge base. There were, however, different opinions about when was the right time to learn about fertility:

When you're 20 years old, it's from there, or at least it's my idea that it's from there that most people think, now we're starting to have an awareness of this, that at some point we're going to have children over the next ten years. (Female, 29 years, Bachelor's education)

Maybe – later during higher education, universities and professional schools and vocational training. But now I am at the age of 18 and have to be scared a little by reduced fertility. I think it's too early. (Male, 18 years, high school)

It's too early for me right now. It's not something I want to decide on right now. (Male, 18 years, technical school)

How, what and from whom

How – the format. There were personal preferences for teaching formats. The participants wanted to hear about fertility in multifaceted ways (e.g. social media, podcasts, theatre, radio). There was no agreement on one single way to address this. There were different responses about the use of campaigns. Some of the participants thought it was provocative to use what they perceived as scare campaigns while others saw it as the only solution to grab their attention:

If you do a scare campaign that shows that this is how far it can go, and this is how you can screw up your life, if you take drugs, etc., then it might make people think, but if you start softly out with 'Would you like to have a child? What considerations have you made?', then I think they are gone. (Male, 28 years, technical school)

Lectures were mentioned by many of the young people as a good solution, while others would prefer being in smaller groups, so that everyone could be involved in the discussion.

When commenting on the poster, many thought it was informative. However, some thought that there was too much text and information and it was too passive. The participants expressed that a poster format was not suitable when addressing their segment of the population (female, 18 years, high school). However, social media was seen as a very important source of information when providing information to young people. They wanted brief information using videos or pictures/images:

Definitely over our phones. I think it's easiest to reach out to us there anyway. (Female, 18 years, high school)

Those little bites, it takes no time, and then you scroll right on. I see everything there because it goes so fast. (Female, 23 years, Bachelor's education)

The youngest participants wanted the information to be provided in a more confrontational manner, whereas the older ones preferred to receive the information in a non-confrontational way. This difference was also apparent between genders, with the men being the ones who preferred the information in a provocative manner. The men also tended to be less worried about their fertility compared with the women.

What – the content. The participants stressed that using personal stories and humour was also important when talking with young people about fertility. Furthermore, honesty was essential.

There was an age divide between the participants on the content of fertility awareness information.

The young adults in this study preferred explanations, rather than listing facts:

Get some more background knowledge so you can understand what is actually happening. Instead, it's just such facts that are thrown into one's head. The fact that you actually get a deeper knowledge of it can create an interest. (Female, 19 years, high school)

During the interviews it became clear that it was important for the educator to see the world through the participants' eyes to get a deeper understanding of them and their life in order for educational efforts to be effective.

From whom – the source. The participants stressed that information should be delivered by, for example, public health or health professionals rather than their daily teachers and that the information should be developed in partnership with the target group (young people). They wanted to feel involved and heard by those who presented the information, and they wanted the possibility to ask the professionals questions:

One's primary school teacher there – just like in the primary school with sex education, where it is one's primary school teacher who thought it was actually a bit awkward. But that thing with some coming from outside. (Female, 22 years, Bachelor's education)

The double-edged sword of knowledge/ consequence of knowledge

Some of the participants described a conflict between the behaviour necessary for a healthy lifestyle to protect their fertility and the behaviour that they believe is a normal aspect of being young:

I think that somewhere I do not really want to know what I really shouldn't do, because I don't want to hear that I should stop drinking alcohol, because it's not going to happen. So the only thing I'm told is that every time I drink alcohol, I might have a little less fun because I think 'Oh no, now I can't have children'. (Male, 19 years, high school)

This participant did not want to know his fertility status or potential risk factors for infertility, because this put him in a dilemma around whether to change his behaviour. Many participants did not know what they were going to do with this new knowledge. The double-edged sword of knowledge and the consequence of knowledge made them hesitant or less open to learning.

Readiness and priorities

The young people expressed that there were many things to think about at this time in their life in terms of their readiness and priorities. Their focus was on their education, friends, love life, appearance and future:

It is difficult to have it all. Both finding the right time to have children, which will probably never exist, and then at the same time wanting to have a future in some subject and get a career and get it off to a good start and focus on your studies. There are many things to juggle. (Male, 19 years, high school)

The participants talked about having freedom without obligations and enjoying life. One of the young women said:

When you are young then you are immortal. (Female, 24 years, Bachelor's education)

All these factors made it difficult for the young people to think about their fertility.

DISCUSSION

In general, the young adults in this study are welcoming the opportunity to learn about fertility. They find it important to learn about fertility from professionals throughout their educational life. While they welcome the information, they also express ambivalence when it comes to the consequences of knowing and the subsequent behaviour. While there are similarities in their opinions and preferences there are also many differences. For education to be effective, it should adopt a multifaceted approach with multiple interventions tailored to the various stages in the lifespan of the audience. This begins with planting the seed during their adolescent years and subsequently expanding the information as they progress.

The young people in the present study demonstrated moderate fertility knowledge scores as measured by the Cardiff Fertility Knowledge Scale, but the results suggest they may not possess the level of knowledge to make informed decisions about their fertility. Fertility knowledge was modest, with a 54% average correct score (13–93%). We consider this to be a modest score relative to the maximum score possible of 100%. Bunting and colleagues found an average score according to the Cardiff Fertility Knowledge Scale of 56.9% among women and men from 79 countries (*Bunting et al., 2013*). Other studies including the Cardiff Fertility Knowledge Scale also found average scores of 51.09% among childless women (*Fulford et al., 2013*), 44.4% among the general population (both women and men; *Maeda et al., 2015*), 49.5% among women in the general population and 42.5% among men in the general population (*Maeda et al., 2016*). Bunting and colleagues found that fertility knowledge is primarily linked to education rather than personal fertility and/or parenting experiences (*Bunting et al., 2013*).

The results of the current study using the Cardiff Fertility Knowledge scale show that average fertility knowledge appears to differ between participants from different educations, with participants from high schools having the highest knowledge and participants from technical schools the lowest; however, this was not significant, which may be due to the small sample size. The young people in the study demonstrated gaps in their fertility

knowledge but also an uncertainty of what they needed to know about fertility. A study by Harper and co-workers found that young people thought that learning about fertility through art was interesting, but they asked for more scientific information and discussion (*Harper et al., 2021*). The current authors tried to include more scientific information through the fertility awareness poster that provides different facts. Furthermore, there was a possibility for discussion among the participants in the focus groups.

This study highlights that age and gender are important factors to consider when developing fertility awareness interventions. The youngest participants and the men wanted the information in a more confrontational way compared with the women and the older participants.

Regardless of age or wish for children, there is a desire for knowledge about fertility. During the study, it was not anticipated that young people who did not want children would have an interest in a fertility awareness study. However, they also want fertility awareness as part of their overall health education or so that they can share this information with others in order to help them achieve their desired reproductive intentions. This was also found by Grace and colleagues, who reported that some of their respondents who felt strongly that it was 'socially irresponsible to have children', for example in the context of climate change or wars, also wanted fertility awareness (*Grace et al., 2022*).

As found in other studies, many people take their fertility for granted and do not wish to think about their fertility until they are trying to conceive a child or until a problem arises with their fertility (*Bodin & Käll, 2020*; *Bodin et al., 2018*; *Hviid Malling et al., 2022*). Thinking actively about one's fertility and having it assessed can lead to existential thoughts and dilemmas about life meaning and legacy. In the current study, some of the participants questioned what they should do with their potentially new knowledge about fertility and were unsure whether it would increase or decrease distress and self-blame and behaviour change. This was also the case in the study by Bodin and collaborators (*Bodin et al., 2021b*). Young people juggle the awareness of fertility decline with the desire to finish education, finding a loving partner, being in phase with their friends and not becoming an aged and tired parent (*Bergnehr, 2008*; *Bodin & Käll, 2020*; *Ragnar et al., 2018*).

During their adolescent years, participants strive to detach from their parents and establish independence. They do not want to be lectured; instead, they desire to embrace youthfulness and freedom. It is therefore crucial to acknowledge and tailor information to the stage of life and perceptions when addressing different segments of the population. This study's results showed a difference between young adults in how to talk about fertility. The youngest were more open to hearing about fertility and risk factors, where the older participants were a bit more reluctant. The older participants may be facing a more urgent dilemma between having knowledge and acting on that knowledge, given they are closer to the end of their reproductive life. For the younger participants there is more time, allowing a separation between gathering knowledge and acting upon that knowledge, which may result in a more open attitude towards receiving information.

It is important to have this in mind when talking with young people about fertility. It would be preferable to use more sensitive language towards people aged 25 years and above. The age at the start of family building has an effect on how many children are desired. In a Swedish report there was a tendency for people who start family building at a younger age to want three or four children, while people starting at a later stage wanted to have only one child (Statistics *Sweden, 2009*). This may reflect on their perception of what is possible within their reproductive years, so they adjust their life and family-building goals to their circumstances, and become more realistic with age. It may also be a way of coping in order to protect themselves from the disappointment of not achieving their family-building goals. It would be preferable if they received the information at an earlier stage in life where they are less sensitive to it. Of course, there could also be other explanations, for example that some people's priorities are different from the very beginning, or that people think having more children is less important than education, career, education or pursuing other goals.

It is also likely that some young people will change fertility intentions throughout their life, as their circumstances change and their family-building goals evolve, and thereby information about fertility can become relevant for them in the future. This needs to be captured, introducing education about fertility at different educational institutions and throughout the lifespan.

Fertility education is also relevant to culture and country. Discussing fertility in an open way is not as possible in all countries as it is in Denmark. Fertility education strategies therefore need to be adapted according to the cultural context. Denmark is part of the Nordic welfare system, which includes day care, free schools, shared parental leave and a heavy involvement in their children's upbringing and well-being. The results of the current study also show that the men have comparable fertility knowledge to the women. Timing of parenthood is a shared decision for most couples in Denmark but this is not the case in all countries. A study by Pearson and colleagues found that Australian men referred either directly or indirectly to fertility as 'women's business' (*Pearson et al., 2021*). This underlines the importance of undertaking studies like this in different settings, cultures and contexts. It is important to co-create educational resources together with the target group (young people) in order to make it relevant.

The strengths and limitations of the study need to be acknowledged. First, the study participants had all chosen to be a part of this study so the results may not be directly transferrable to all young people in terms of attitudes towards how to talk to young adults about fertility. Another weakness is the fact that because recruitment was through educational institutions, the sample failed to include those young adults who discontinued formal education after completing their compulsory schooling. As a result, it could be assumed that the study participants were, at least to some degree and for various reasons, willing to be educated. However, it was possible to interview young people from different parts of Denmark and from different types of educational institution, which is a major strength of this study. The majority of those agreeing to participate in the interview survey were women despite the authors' best efforts to recruit men. The gendered nature of interest in fertility may reflect social norms and/or a lower male interest in fertility.

CONCLUSIONS

This study contributes to the understanding and implementation of

future interventions and campaigns for fertility awareness education targeted to and effective among young adults. These findings can be useful in the process of increasing the fertility awareness in this population.

Improving fertility education in schools and providing all young adults with a foundation of evidence-based knowledge seems to be the best way forward. The participants' preferences for learning about fertility and their reactions to fertility information vary widely on an individual basis. Consequently, information should be customized for the audience and delivered at an appropriate stage in their lifespan, beginning with introducing key concepts when they are young.

DATA AVAILABILITY

The data that has been used is confidential.

ACKNOWLEDGEMENTS

The authors would like to acknowledge all participants for their time and effort.

FUNDING

This study is a ReproUnion collaborative study, co-financed by EU Interreg, Capital Region of Denmark, Region Skåne and Ferring Pharmaceuticals.

REFERENCES

Aguinaldo, ET, Morgan, DC, Julliard, K., 2014. What would you do if you knew? Obstetrics and Gynecology 123, 187S.

Bergnehr, D., 2008. Timing parenthood: Independence, family and ideals of life. Department of Child Studies, Linköping University. http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-11440.

Berthelsen, ASN, Gamby, ALN, Christensen, U, Schmidt, L, Koert, E., 2021. How do young men want to receive information about fertility? Young men's attitudes towards a fertility campaign targeting men in Copenhagen, Denmark. Human Reproduction Open hoab027. https://doi.org/10.1093/hropen/hoab027.

Bodin, M, Käll, L., 2020. Is it an issue before it's a problem? Investigating men's talk about fertility. Sociology of Health & Illness 42, 1611–1625. https://doi.org/10.1111/1467-9566.13148.

Bodin, M, Holmström, C, Plantin, L, Schmidt, L, Ziebe, S, Elmerstig, E., 2021a. Preconditions to parenthood: changes over time and generations. Reproductive Biomedicine & Society Online 13, 14–23.

Bodin, M, Plantin, L, Schmidt, L, Ziebe, S, Elmerstig, E., 2021b. The pros and cons of fertility awareness and information: a generational, Swedish perspective. Human Fertility. https://doi.org/10.1080/14647273.2021.1968045.

Bodin, M, Tyden, T, Käll, L, Larsson, M., 2018. Can Reproductive Life Plan-based counselling increase men's fertility awareness? Upsala Journal of Medical Sciences 123, 255–263. https://doi.org/10.1080/03009734.2018.1541948.

Bowen, DJ, Kreuter, M, Spring, B, Cofta-Woerpel, L, Linnan, L, Weiner, D, Bakken, S, Patrick Kaplan, C., Squiers, L, Fabrizio, C, Fernandez, M, 2009. How we design feasibility studies. American Journal of Preventive Medicine 36, 452–457. https://doi.org/10.1016/j.amepre.2009.02.002.

Bunting, L, Tsibulsky, I, Boivin, J., 2013. Fertility knowledge and beliefs about fertility treatment: findings from the International Fertility Decision-making Study. Human Reproduction 28, 385–397 . https://doi.org/10.1093/humrep/des402.

Brandt, JS, Cruz Ithier, MA, Rosen, T, Ashkinadze, E, 2019. Advanced paternal age, infertility, and reproductive risks: A review of the literature. Prenatal Diagnosis 39, 81–87. https://doi.org/10.1002/pd.5402.

Cox, CM, Thoma, ME, Tchangalova, N, Mburu, G, Bornstein, MJ, Johnson, CL, Kiarie, J., 2022. Infertility prevalence and the methods of estimation from 1990 to 2021: a systematic review and meta-analysis. Human Reproduction Open 1–24. https://doi.org/10.1093/hropen/hoac051.

Daniluk, JC, Koert, E., 2013. The other side of the fertility coin: A comparison of childless men's and women's knowledge of fertility and assisted reproductive technology. Fertility and Sterility 99, 839–846. https://doi.org/10.1016/j.fertnstert.2012.10.033.

Delbaere, I, Pitsillos, T, The Greek Collaborating Group, Tydén, T, Kerckhof, L, Iliadis, SI, 2021. Fertility awareness and parenthood intentions among medical students in three European countries. The European Journal of Contraception & Reproductive Health Care 26 (4), 312–322. https://doi.org/10.1080/13625187.2021.1901877.

Fulford, B, Bunting, L, Tsibulsky, I, Boivin, J., 2013. The role of knowledge and perceived susceptibility in intentions to optimize fertility: findings from the International Fertility Decision-Making Study (IFDMS). Human Reproduction 28, 3253–3262.

Garcia, D, Vassena, R, Prat, A, Vernaeve, V., 2016. Increasing fertility knowledge and awareness by tailored education: a randomized controlled trial. Reproductive Biomedicine Online 32, 113–120. https://doi.org/10.1016/j.rbmo.2015.10.008.

Grace, B, Shawe, J, Johnson, S, Usman, NO, Stephenson, J., 2022. The ABC of reproductive intentions: a mixed-methods study exploring the spectrum of attitudes towards family building. Human Reproduction 37, 988–996. https://doi.org/10.1093/humrep/deac036.

Graneheim, UH, Lundman, B., 2004. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Education Today 24, 105–112. https://doi.org/10.1016/j.nedt.2003.10.001.

Hammarberg, K, Collins, V, Holden, C, Young, K, McLachlan, R., 2017. Men's knowledge, attitudes and behaviours relating to fertility. Human Reproduction Update 23, 458–480.

Harper, JC, Hepburn, J, Vautier, G, Callander, E, Glasgow, T, Balen, A, Boivin, J., 2021. Feasibility and acceptability of theatrical and visual art to deliver fertility education to young adults. Human Fertility 24, 129–135. https://doi.org/10.1080/14647273.2019.1570354.

Hviid Malling, GM, Schmidt, L, Pitsillos, T, Hammarberg, K, Tyden, T, Friberg, B, Jensen, I, Ziebe, S, 2022. Taking fertility for granted – a qualitative exploration of fertility awareness among young, childless men in Denmark and Sweden. Human Fertility 25, 337–348. https://doi.org/10.1080/14647273.2020.1798516.

Maeda, E, Nakamura, F, Kobayashi, Y, Boivin, J, Sugimori, H, Murata, K, et al., 2016. Effects of fertility education on knowledge, desires and anxiety among the reproductive-aged population: findings from a randomized controlled trial. Human Reproduction 31, 2051–2060.

Maeda, E, Sugimori, H, Nakamura, F, Kobayashi, Y, Green, J, Suka, M, Okamoto, M, Boivin, J, Saito, H, 2015. A cross sectional study on fertility knowledge in Japan, measured with the Japanese version of Cardiff Fertility Knowledge Scale (CFKS-J). Reproduction Health 12, 1–12.

Pearson, L, Holton, S, McLachlan, R, Hammarberg, K., 2021. Australian men's fertility information seeking attitudes and behaviour: A qualitative investigation. Sexual & Reproductive Healthcare 29, 100621.

Pedro, J, Brandão, T, Schmidt, L, Costa, ME, Martins, MV., 2018. What do people know about fertility? A systematic review on fertility awareness and its associated factors. Upsala Journal of Medical Sciences 123, 71–81.

Ragnar, ME, Grandahl, M, Stern, J, Mattebo, M., 2018. Important but far away: adolescents' beliefs, awareness and experiences of fertility and preconception health. The European Journal of Contraception & Reproductive Health Care: The Official Journal of the European Society of Contraception 23, 265–273. https://doi.org/10.1080/13625187.2018.1481942.

Ren, Y, Xie, Y, Xu, Q, Long, M, Zheng, Y, Li, L, Niu, C, 2023. University students' fertility awareness and its influencing factors: a systematic review. Reproductive Health 20, 85.

Saunders, B, Sim, J, Kingstone, T, Baker, S, Waterfield, J, Bartlam, B, Burroughs, H, Jinks, C., 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. Quality & quantity 52, 1893–1907.

Schmidt, L, Münster, K, Helm, P., 1995. Infertility and the seeking of in fertility treatment in a representative population. British Journal of Obstetrics and Gynaecology 102, 978–984.
Statistics Denmark, 2023. Vital Statistics, 2023. Statistics Denmark, Copenhagen.
Sweden, Statistics, 2009. Barn eller inte? Statistics Sweden, Örebro.
Tong, A, Sainsbury, P, Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care 19, 349–357. https://doi.org/10.1093/intqhc/mzm042.
Vassard, D, Lallement, C, Nyboe Andersen, A, Macklon, N, Schmidt, L, 2016. A population-based survey on family intentions and fertility awareness in women and men in the United Kingdom and Denmark. Upsala Journal of Medical Sciences 121, 244–251.
Zegers-Hochschild, F, Adamson, GD, Dyer, S, Racowsky, C, Mouzon, J, Sokol, R, Rienzi, L, Sunde, A, Schmidt, L, Cooke, ID, Simpson, JL, van der Poel, S., 2017. The International Glossary on infertility and fertility care. Human Reproduction 32, 1786–1801.

Received 4 October 2023; received in revised form 4 March 2024; accepted 4 March 2024.